l'Université Lille 1 - Sciences et Technologies,
Ecole doctorale SESAME, Laboratoire EQUIPPE
l'Université Fédéral de l'Oural

Thèse

Pour obtenir le grade de

Docteur

Spécialité: Sciences Economiques

Présentée et soutenue par

Natalia DAVIDSON

le 9 décembre 2013 à l'Université Lille 1 - Sciences et Technologies

The impact of spatial concentration on enterprise performance and location choice in Russia

Thèse dirigée par:

M. Hubert JAYET, Professeur à l'Université Lille 1 - Sciences et Technologies /
M. Sergey KADOCHNIKOV, Professeur à l'Université Fédéral de l'Oural

Jury:

M. Glenn RAYP, Professeur à l'Université de Gand	Rapporteur
M. Oleksandr SHEPOTYLO, Professeur à l'Ecole Supérieure d'Economie, Moscou	Rapporteur
M. Sergey AFONTSEV, Professeur à l'Université Fédéral de l'Oural ; Institut d'Economie Mondiale et des Relations Internationales, l'Académie des Sciences, Moscou	Examinateur
Mme Lise PATUREAU, Professeur à l'Université Lille 1 – Sciences et Technologies	Examinateur

Titre : L'impact de la concentration spatiale sur la performance des entreprises en Russie et leur choix de localisation

Résumé

Les données sur les firmes, les villes et les régions russes des années 1999-2008 sont utilisées pour analyser l'impact de l'agglomération et du potentiel de marché (PM) sur la productivité des firmes et leur choix de localisation. Le salaire, l'infrastructure de transport et le climat économique sont pris en compte. La motivation est de trouver les sources de développement des villes, dont les conditions initiales et la situation actuelle sont considérablement différentes. Les économies d'agglomération sont basées sur le partage, l'adéquation et l'apprentissage. Selon le Chapitre 3, les économies d'urbanisation et de diversité ainsi que l'effet de PM sont positifs. Les économies de localisation ont la forme de U inversé dans la plupart des cas, mais elles sont positives pour les monovilles ; les économies d'urbanisation sont aussi importantes pour celles-ci. Le Chapitre 4 montre que les économies d'urbanisation résultant de la présence des firmes nationales et étrangères sont positives ; en même temps, les économies générées par les firmes étrangères sont plus importantes. Les économies de localisation qui proviennent des firmes nationales ont la forme de U inversé ; celles qui proviennent des firmes étrangères sont contradictoires. Le Chapitre 5 étudie le choix de localisation par les firmes parmi les villes. Les économies d'urbanisation ainsi que le PM ont l'impact positif sur un choix de localisation. Les économies de localisation gardent la forme de U inversé. Le choix de localisation par les firmes étrangères est plus influencé par l'agglomération, le PM et le climat économique que celui fait par les firmes nationales à cause des possibilités de choix plus grandes.

Les mots clés : économies d'agglomération, potentiel de marché, investissement direct étranger, choix de localisation des entreprises, villes, Russie

Abstract

Firm, city and regional level data for Russia, years 1999-2008, is employed to analyze the effects of agglomeration level and home market potential (HMP) on enterprise productivity and location choice. City average wages, transport infrastructure and business environment are considered. Our motivation is search for sources of economic development in the Russian cities, which differ significantly in their initial conditions and present situation. Agglomeration economies are associated with the opportunities for input sharing, matching and knowledge spillovers. In Chapter 3, urbanization and diversity economies as well as HMP are found to be positive. Localization economies have an inverted U shape for the majority of specifications, but are positive for the monotowns, showing the importance of firms' concentration in the same industry for enterprise performance. Urbanization economies are important for the monotowns too. In Chapter 4, it is found that urbanization economies arising from presence both of national and foreign firms are positive, while those arising from the foreign firms' presence are relatively larger. Localization economies arising from the national firms' presence have an inverted U shape; localization effects arising from the foreign firms' presence are contradictory. In Chapter 5, enterprise location choice for a city is estimated. Both urbanization economies and HMP positively affect location choice; localization economies keep the inverted U shape. Foreign firms' choice for a city is more affected by agglomeration, HMP and institutional infrastructure, probably as they have greater opportunities than the national firms to choose location.

Key words and phrases: agglomeration economies, home market potential, foreign direct investment, enterprise location choice, cities, Russia

Résumé détaillé

Comme dans les autres pays, en Russie, les régions et les villes différent concernant leur situation économique. Néanmoins, le but de la science economique est de contribuer au niveau de vie élevé pour chacun dans le pays ayant les resources limitées. Les opportunités de développement diffèrent entre villes et régions dans la concurrence. Dans l'Union Sovétique les possibilités de changement de lieu d'habitation étaient moindres. La transition vers l'économie de marché a donné aux gens et aux entreprises quelques opportunitées à choisir entre les villes. La diversité, ouverture, échange d'idées renforcent l'attrait des territoires et leur dévelopment économique.

Dans notre contexte on comprend la diversité comme la variété des activités économiques dans les villes ; ce terme comprend aussi les effets de l'investissement direct étranger (IDE) sur la performence des entreprises et leur choix de localisation. Selon la théorie du cycle de vie (Vernon, 1960), les nouvelles branches de l'économie bénéficient de la localisation dans des villes avec des activités économiques diversifiées. L'IDE favorise le développement économique, la création de nouveaux emplois dans les pays-hôtes et l'efficacité des chaînes de production. Cependant, les coûts de la congestion caractérisent les grandes villes ; l'IDE est associé aux effets controversés résultant de la concurrence entre les entreprises étrangères plus avancées et les firmes locales. C'est pourquoi ces questions doivent être étudiée plus en détail.

L'object de cette thèse est de donner un aperçu sur l'impact de la concentration spatiale sur la performance des entreprises en Russie et leur choix de localisation. La concentration spatiale comprend la localisation, la diversité et l'urbanisation, et le potentiel de marché de la ville. On considère les mécanismes sur lesquels sont basées les économies d'agglomération. En outre, le capital humain, l'infrastructure de transport et les risques de la conjoncture économique au niveau régional sont pris en compte. Plusieurs indices d'agglomération sont analyses, et une version d'un indice de la diversité proposé par Vorobyev et al. (2010) est utilisée.

L'objectif des Chapitres 3 et 4 est d'estimer l'impact du niveau d'agglomération et du potentiel de marché sur la productivité des entreprises dans une ville. Dans le Chapitre 4, les effets d'agglomération sont subdivisés en ceux provenant de la concentration des firmes nationales et étrangères. Particulièrement, il est important d'éclairer l'impact de la concentration des firmes étrangères dans une ville sur la performance des firmes nationales. L'objectif du Chapitre 5 est de trouver les facteurs déterminant le choix de localisation par les firmes entre les villes. La probabilité 'du choix de localisation' peut être interprétée pour les firmes étrangères comme, en fait, la probabilité de localisation dans une ville, et pour les

firmes nationales comme la probabilité de fondation et de fonctionnement dans une ville. Les facteurs déterminant le choix de localisation des enterprises avec les caractéristiques différentes sont analysés. La situation dans des types de villes différentes, y compris les monovilles, est examinée.

La recherche actuelle sur les économies d'agglomération. Actuellement, l'économie spatiale attire l'intérêt en tant que la Novelle Géographie Economique, commençant par les oeuvres de Krugman (1991), Fujita (1988) et Venables (1996). Krugman (1991) remarque que le schéma de l'activité économique en Europe est reflété par les photos satellite de nuit indiquant la concentration de l'activité en Belgique ou aux environs. Marshall (1920) a supposé que, parmi les raisons qui sont à l'origine de la concentration spatiale des secteurs industriels, il y a 'les mystères du métier'.

Les effets positifs et négatifs d'échelle et d'étendue*(scope)* des activités qui résultent de la concentration de l'activité économique sont définis comme les externalités d'agglomération (Neffke, 2009). Ils proviennent de la concentration des firmes similaires *(localisation ou spécialisation – les externalités de Marshall – Arrow – Romer (MAR))* ou des firmes qui appartiennent aux industries variées *(diversité ou urbanisation ; les externalités de Jacobs).* La diversité industrielle est importante pour la sécurité sociale et pour la croissance économique urbaine (Jacobs, 1961). Selon le concept du potentiel de marché (Home market potential) dans La Nouvelle Géographie Economique, si une économie comprend deux régions et si les secteurs sont caractérisés par la concurrence imparfaite, dans ce cas la part plus que proportionnelle des firmes choisissent la région avec une plus grande demande (Ottaviano, Thisse, 2004).

Les économies d'agglomération sont basées sur 3 mécanismes principaux. Premièrement, il y a le partage des ressources indivisibles (*sharing)*. Deuxièmement, les agglomérations sont caractérisées par l'adéquation – l'existence du marché de travail spécialisé qui permet un meilleur appariement entre les travailleurs et les employeurs (*matching*). Le troisième mécanisme est l'apprentissage facilité par la circulation de connaissances entre les gens et les entreprises dans les milieux avec des activités diversifiées (*learning*) (Duranton, Puga, 2004).

La localisation des firmes à la proximité immédiate des unes des autres reste pertinente à l'époque des technologies de télécommunications, partiellement en raison de la connaissance tacite (Audretsch, 1998; Robert-Nicaud, 2013). D'autre côté, Mills (1967) indique la possibilité de la congestion dans une ville causée particulièrement par les coûts de transport et de logement, c'est-à-dire les économies d'agglomération sont supposées avoir la forme de U inversé. Les résultats concernant les effets d'agglomération sont variés (Beaudry

et Schiffauerova, 2009; De Groot et al. 2009). L'estimation des effets d'agglomération dépend du contexte (Neffke, 2009). La prime de productivité causée par le doublement de la taille d'une ville aux Etats-Unis est 3 – 8% (Rosenthal et Strange, 2004).

Henderson et al. (1995) trouvent une preuve que des externalités de localisation existent pour les industries matures, et en ce qui concerne les industries nouvelles il y a des externalités de localisation ainsi que celles de diversité. Les effets de diversité tendent à être significatifs pour les nouvelles entreprises, tandis que les économies de localisation sont plus importantes pour les entreprises matures (Duranton and Puga, 2001; Neffke et al., 2008).

Baldwin et Okubo (2005) introduisent l'approche de Melitz (2003) avec les firmes hétérogènes dans le modèle de la Nouvelle Géographie Economique et révèlent que l'hétérogénéité des firmes réduit les effets du potentiel de marché, et que les firmes plus efficaces sont attirées par les régions plus grandes. Vakhitov et Bollinger (2010) utilisent les données sur l'Ukraine et trouvent les effets de localisation et ceux de diversité (d'urbanisation) dans la fabrication des machines ainsi que dans les secteurs des hautes technologies ; les économies d'agglomération obtenues par les firmes étrangères sont les plus grandes.

Pour la Russie, la taille et l'étendue *(scope)* des activités des villes ainsi que les retombées de la croissance entre les régions sont les plus importantes (Russian Economic Report, 2007). L'Institut d'Analyse des Entreprises et du Marché, L'Ecole Supérieure d'Economie (IAEM, *Higher School of Economics*), en coopération avec la Banque mondiale, le FMI et le 'Lévada-Centre' a étudié les entreprises russes des industries manufacturières et a révélé que le doublement de la taille d'une ville mène à l'accroissement de la productivité de 5%. On est arrivé à la conclusion que les firmes sont moins concurrencielles dans les petites villes et les monovilles, mais elles peuvent être concurrencielles s'il y a là assez d'autres entreprises, y compris les petites entreprises (Enterprises and markets, 2010).

Gonchar (2010) utilise les données pour la Russie des années 2005-2006 et trouve que les entreprises qui appartiennent aux agglomérations sont en moyenne de 46% plus productives que les autres entreprises, c'est-à-dire, les effets d'échelle externes sont positifs. En même temps, en Russie les économies d'échelle internes dominent, la productivité des firmes étant influencée par leur taille (Gonchar, 2010). Vorobyev et al. (2010) analysent les économies de localisation et de diversité pour les villes russes se basant sur les données pour les firmes manufacturières des années 2001-2004. Ils trouvent les économies de localisation caractérisées par la forme de U inversé et les économies de diversité positives dans les villes.

La recherche actuelle sur l'investissement direct étranger (IDE). L'IDE est l'investissement en vue d'organiser une filiale dans un pays étranger ou d'exercer le contrôle

sur la gestion d'une entreprise existant à l'étranger. Les effets de l'IDE sont les effets directs et les retombées de productivité (les retombées horizontales, verticales et '*supply-backward*') (Smarzynska-Javorcik, 2004; Merlevede, Schoors, 2008). La possibilité de transfert technologique est parmi les raisons pour lesquelles les gouvernements visent à attirer l'IDE (Blomström and Kokko, 2003).

Les firmes étrangères essaient de protéger leurs informations et technologies. Cependant, elles sont intéressées par la haute qualité des produits intermédiaires. Par conséquent, les retombées verticales de l'IDE sont plus probables d'être positives que les retombées horizontales (Smarzynska-Javorcik, 2004). Dans le cadre des économies d'agglomération, les retombées horizontales sont associées avec les effets de localisation (Blomström and Kokko, 2003). Les effets verticaux peuvent être liés aux effets de diversité ou d'urbanisation.

Les modèles avec les firmes hétérogènes sont appliqués dans le but d'analyser les questions concernant l'IDE. Ils prennent en considération les faits suivants. Premièrement, la productivité des entreprises dans un secteur a tendance à différer substantiellement. Deuxièmement, les firmes plus productives sont normalement plus engagées dans le commerce international et dans l'IDE. Troisièmement, la libéralisation du commerce conduit à la redistribution des parts de marché entre les firmes dans un secteur, les firmes plus productives obtenant les parts plus grandes. Finalement, les changements dans les parts de marché parmi les firmes d'une industrie provoquent la croissance de la productivité moyenne dans l'industrie, même si la productivité de chaque firme reste constante (Melitz, 2003).

La preuve empirique sur les effets de l'IDE montre que les firmes dans les pays développés ont les retombées de l'IDE plus importantes, probablement en raison de leur capacité d'absorption plus élévée (Aitken and Harrison, 1999; Smarzynska-Javorcik, 2004; Bode et al., 2009). Crozet et al. (2004) fournissent la preuve que l'attraction de l'IDE dans une industrie peut augmenter la productivité dans cette industrie et mener au changement de l'avantage comparatif d'un pays ou d'une région. Bode et al. (2009), se basant sur les données pour les Etats-Unis des années 1977-2003, ont révélé que les firmes avec l'IDE sont une source des externalités de localisation positives, tandis que les firmes domestiques sont une source des externalités négatives.

La taille et le signe des effets de l'IDE varient, dépendant de la distance géographique entre la filiale étrangère et la firme nationale, de la période du transfert technologique, des caractéristiques institutionnelles et politiques du pays hôte, des caractéristiques d'une industrie, des stratégies d'une corporation multinationale, de la taille, de la productivité et de la capacité d'absorption des firmes nationales, de l'écart technologique (Borenzstein et al.

1998; Hall and Jones, 1999; Bekes, Kleinert, Toubal, 2009; Altomonte and Resmini, 2001; Damijan et al., 2013).

La recherche utilisant les données du debut de la transition, montre plus de preuves des retombées verticales positives que de preuves des retombées horizontales positives (Altomonte and Resmini, 2001; Merlevede and Schoors, 2008; Elmas and Degirmen, 2009; Smarzynska-Javorcik, 2004; Stancik, 2007). Damijan et al. (2013) trouvent les retombées horizontales positives pour les pays en transition, qui deviennent plus importantes pendant les dix dernières années. Selon eux, les retombées horizontales négatives sont typiques pour les plus petites firmes, et pour les firmes avec un bas ou moyen niveau de productivité.

Pour la Russie, les retombées positives ainsi que négatives de l'IDE sont révélés (Yudaeva et al., 2003; Kadochnikov et Drapkin, 2007; Drapkin et al.,2011a). Dolgopyatova (2008) conclut que l'IDE entraîne le développement des pratiques du business en Russie au niveau des firmes et influence la politique du gouvernement. Drapkin et al. (2011a) révèlent les effets horizontaux positifs dans les régions russes avec la structure économique diversifiée, et les effets négatifs dans les régions spécialisées pour les années 1999-2008. Nous montrons que dans les centres d'agglomération (les villes diversifiés), les effets de localisation provenant des firmes étrangères sont positifs pour les firmes nationales. Dans les autres villes ils sont insignificatifs. En moyenne, pour les firmes nationales produisant les biens marchands les effets de localisation provenant des firmes étrangères sont négatifs.

L'accession de la Russie à l'Organisation Mondiale du Commerce est l'un des faits centraux influençant le rôle du pays dans le commerce international et la situation économique nationale. Rutherford et Tarr (2008) développent un modèle d'équilibre général calculable et utilisent les données du Rosstat sur les ménages. Ils concluent que les facteurs cruciaux sont l'IDE dans le secteur des services aux entreprises et les effets endogènes de la productivité générés par les variétés supplémentaires des services aux entreprises et des biens. Selon ce modèle, les gains du bien-être national sont environs 4,5% du PIB.

La recherche actuelle sur le choix de localisation par les firmes. Head et Mayer (2004) combinent les concepts de la taille du marché et des économies d'agglomération. Ils étudient le potentiel de marché en utilisant un modèle théorique du choix de localisation qui donne une fonction du profit. Cette fonction contient l'indice du potentiel de marché introduit par Harris (1954), la distribution spatiale des concurrents et les variables d'agglomération. Se basant sur les données japonaises au niveau des enterprises des années 1984-1995, Head et Mayer (2004) trouvent que le choix de localisation est enfluencé par le potentiel de marché et par les forces d'agglomération.

Selon les résultats de Crozet et al. (2004), la politique regionale ne contribue pas de manière significative à l'attraction de l'IDE. Néanmoins, l'expérience de certains pays, par exemple, l'Irelande, démontre que la politique économique influence les projets avec l'IDE (Navaretti, Venables, 2004; Barry et al. (2001). Afin d'attirer des entreprises, le climat des investissments et l'image favorable auprès des investisseurs sont essentiels, le climat économique étant composé d'aspects économiques, politiques et sociaux (Shastitko et al., 2007).

Le choix de localisation des firmes étrangères est lié aux stratégies des firmes multinationales concernant l'IDE. Ce sont les stratégies traditionnelles (stratégie d'approvisionnement, stratégie de marché, recherche d'actifs, stratégie de rationalisation) et les stratégies globales plus complexes (Dunning, 1993, Andreff, 2003). D'après les résultats reçus par Melitz (2003), seulement les firmes les plus productives entrent dans le marché étranger. Helpman et al. (2004) trouvent aussi que les industries caractérisées par l'hétérogénéité des firmes plus importante réalisent plus de ventes par le biais de l'IDE que par le biais de l'export. Helpman et al. (2004) confirment leurs conclusions théoriques se basant sur les données pour les Etats-Unis de l'année 1994.

Boudier-Bensebaa (2005) utilise les données des années 1991-2000 pour la Hongrie et trouve que la disponibilité de main-d'oeuvre, la demande industrielle, la densité manufacturière, les économies d'agglomération inter-industrielles et l'infrastructure sont significatives pour attirer l'IDE ; les firmes étrangères préfèrent les régions avec les coûts du travail élevés. Belitsky (2011) utilise les données sur les villes de l'ancienne Union soviétique des années 1995-2008 pour estimer la variation du taux de croissance du PIB par habitant et trouve que les facteurs influençant la croissance économique des villes sont suivants : la taille de l'administration locale, le niveau de pauvreté et la qualité des institutions, y compris la protection des droits de la propriété. Shepotilo (2011) étudie les pays d'Europe et d'Asie Centrale et conclut que dans l'Union soviétique la distribution de la population était plus équilibrée qu'elle n'aurait été dans l'économie de marché. Maintenant la mobilité de la population reste basse, et la convergence vers le nouvel equilibre spatial est basse à cause de l'infrastructure des villes insuffisamment développée et de la politique du gouvernement non-cohérente.

Ledyaeva (2009) trouve que les déterminants significatifs de l'IDE pendant les années 1995-2006 pour la Russie sont suivants : la présence du pétrole et du gaz, les avantages d'agglomérations, le niveau du développement industriel, la taille de marché et les avantages des grandes villes, les avantages des régions de transit et la protection des investisseurs par la législation locale. Elle constate la concentration régionale forte et croissante de l'IDE dans

l'économie russe. Ledyaeva (2010) analyse le choix de localisation des firmes chinoises, japonaises et américaines en Russie, utilisant les données au niveau des entreprieses et confirme que la plupart de l'IDE est basé sur la stratégie de marché. Elle trouve que les déterminants importants de localisation de l'IDE dans les régions sont le potentiel de marché, l'infrastructure de transport, développement de la démocratie et les effets d'agglomération.

En utilisant le modèle logit conditionnel et les données pour les firmes russes avec l'IDE ainsi que les données régionales des années 2000-2009, Gonchar et Marek (2013) révèlent que le schéma géographique de choix de localisation de l'IDE en Russie est déterminé par la disponibilité des ressources naturelles et par les avantages liés au marché. Ils trouvent aussi que les effets d'agglomération proviennent de la concentration des firmes d'une même industrie ainsi que des industries différentes ; les investissements dans les services aux enreprises sont co-localisés avec l'IDE lié aux ressources naturelles. Castiglione et al. (2012) étudient les facteurs d'attraction de l'IDE dans les régions russes, en utilisant les données des années 1996 – 2006. Ils montrent que la distribution de l'IDE très inégale est influencée par les facteurs classiques de la demande et par les caractéristiques régionales économiques et socio-institutionnelles spécifiques. Castiglione et al. (2012) concluent aussi que la distribution de l'IDE pendant la période entre 1996-2001 a influencé l'entrée d'IDE pendant 2002-2006.

La taille du marché, la population, les coûts de transport et la distance sont les déterminants 'classiques' de la localisation de l'IDE . Cependent, pour les pays en transition à l'environnement institutionnel faible, les caractéristiques sociales et institutionnelles sont importantes (Castiglione et al., 2012; Boudier-Bensebaa, 2005; Ledyaeva et al., 2010).

Méthodologies. La méthodologie est décrite dans le Chapitre 2. Les données en panel des années 1999-2008 sont utilisées pour évaluer l'impact de l'agglomération et du potentiel de marché sur la productivité des firmes et leur choix de localisation. La recherche utlise les données des enterprises, des villes et des régions russes. A l'instar de la recherche actuelle, l'agglomération est subdivisée en localisation et urbanisation (ou la diversité). Le modèle est estimé au niveau des villes. Le salaire, l'infrastructure de transport et le climat économique au niveau régional sont pris en compte.

Afin de faire l'analyse plus précise, nous faisons la distinction entre les groupes d'industries, les types de villes, et les types d'entreprises. En outre, les effets fixes d' entreprise et les effets fixes temporels sont introduits. *Les industries* sont subdivisées en cinq groupes : les industries traditionnelles, les matériaux de base, la fabrication des machines de base, la fabrication intégrante des machines, et les industries de haute technologie. Quant à la classification des territoires, *les villes* sont classifiées dans les types suivants. Premièrement, les centres des agglomérations - les villes avec la population plus de 1 mn et/ou avec l'activité

économique élevée. Deuxièmement, les villes appartenant aux agglomérations, ces dernières étant définies comme territoires dans le rayon de 60 km autour des centres d'agglomérations. Troisièmement, les monovilles – les villes spécialisées, où l'économie est dominée par une seule industrie. Finalement, toutes les autres villes. Les agglomérations matures en Russie sont localisées principalement dans la partie Européenne centrale du pays, dans la région de Volga, dans l'Oural, et le long du Chemin de fer Transsibérien (Gonchar, 2008).

Dans le Chapitre 3, l'impact de la concentration spatiale sur la performance des entreprises est évalué. Les effets fixes sont utilisés pour régler l'endogénéité provenant de l'hétérogénéité inaperçue entre les firmes. L'endogénéité provenant de la simultanéité est supposée être absente pour les raisons expliquées dans les Chapitres 2 et 3. Dans le Chapitre 4, la concentration spatiale est subdivisée en la concentration des firmes nationales et étrangères. La première différenciation est utilisée afin de traiter l'endogénéité provenant de l'hétérogénéité inobservée entre les firmes. Les variables instrumentales (les variables d'agglomération retardées de deux périodes) et la Méthode des moments généralisées (GMM) sont utilisés afin de traiter l'endogénéité causée par la simultanéité.

Dans le Chapitre 5, en utilisant les données de l'année 2007, le modèle logit conditionnel est estimé. Les raisons déterminant le choix de localisation des enterprises sont examinées principalement à la base des externalités d'agglomération et du potentiel de marché. Les décisions de localisation des firmes nationales et étrangères sont étudiées. Comme dans les chapitres précédents, les institutions et l'infrastructure régionale sont prises en compte. L'estimation est basée essentiellement sur l'équation des bénéfices proposée par Head et Mayer (2004). On suppose que les firmes choisissent les villes où elles peuvent gainer les profits les plus élevés. Le modèle logit conditionnel est utilisé pour estimer les paramètres de l'équation des bénéfices, c'est-à-dire, le choix de localisation des entreprises (MacFadden, 1974, 1978; Head and Mayer, 2004; Combes et al., 2008a). La variable dépendente est le choix de localisation dans une ville. Le modèle est estimé pour les firmes nationales et pour les firmes étrangères (Combes et al. (2008a)).

Les résultats principaux. Le Chapitre 3 confirme l'hypothèse que les économies d'urbanisation et de diversité sont positives. Les économies de localisation ont la forme de U inversé dans la plupart des cas, autrement dit, les économies d'agglomération décroissent à cause de la congestion après qu'un certain niveau d'activité économique soit atteint. L'effet du potentiel de marché, provenant des autres villes, est significatif et positif dans la plupart des spécifications. Vu que cet effet est lié à la distance et aux frais de transport, l'amélioration de l'infrastructure de transport et la diminution des coûts de communication pourraient faciliter l'accroissement de cet effet.

Les effets d'agglomération varient parmi les industries. Pour les firmes privées ces effets sont plus forts qu'en moyenne pour les firmes produisant les biens marchands ; pour les firmes étrangères ces effets sont encore plus importants. Lorsque les types des villes sont concernés, l'impact de localisation et d'urbanisation sur la performance des entreprises est le plus élévé pour les monovilles. L'effet du potentiel de marché est important principalement pour les villes à l'extérieur des agglomérations.

Quant au niveau de localisation optimal, les situations pour les entreprises pourraient être subdivisées en trois types. Premièrement, il y a des firmes localisées dans les petites villes avec la faible concurrence et le manque d'opportunités de développement. La politique raisonable pour ces villes serait d'améliorer le climat économique afin d'attirer d'autres entreprises. Ce serait convenable aussi de développer l'infrastructure de transport autour de ces villes, en vu d'augmenter leur interaction avec les territoires voisins.

Deuxièmement, il y a des firmes localisées près du niveau optimal de localisation qui intériorisent les économies de localisation. Finalement, il y a les villes où les firmes font face aux effets de localisation décroissants. Pour elles, l'ouverture au commerce et à la migration semble pertinente, dans la situation de la concurrence pour les ressources et les consommateurs dans une ville. Les possibilités d'ouverture au commerce dépendent à leur tour de l'amélioration des réglementations officielles et du climat économique, du développement des services aux entreprises et de l'infrastructure de transport.

Le Chapitre 4 montre que les économies d'urbanisation résultant de la présence des firmes nationales et étrangères sont positives ; en même temps, les économies générées par les firmes étrangères sont plus importantes dans le modèle avec la première différenciation et le GMM. Ces effets caractérisent les firmes nationales et les firmes étrangères, tandis que pour les firmes étrangères les effets de la presence des autres firmes étrangères dans les industries variées dans une ville sont plus élevés que ceux pour les firmes nationales. Les effets d'urbanisation élevés et robustes peuvent être expliqués par les opportunités d'apprentissage, de partage et, jusqu'à un certain degré, par l'adéquation dans des grandes villes.

Les économies de localisation qui proviennent des firmes nationales ont la forme de U inversé dans la plupart des cas. Bien que ces économies diminuent après qu'un certain niveau d'activité économique soit atteint, elles restent positives pour les niveaux de localisation actuels dans les échantillons. Les économies de localisation qui proviennent des firmes étrangères sont contradictoires, étant positives pour les firmes situées dans les centres des agglomérations. L'effet du potentiel de marché est positif.

Le Chapitre 5 étudie le choix de localisation par les firmes entre les villes. Les variables explicatives similaires à celles dans les Chapitres 3 et 4 sont considérées. Les

résultats confirment l'hypothèse selon laquelle le niveau d'agglomération et le potentiel de marché influencent la probabilité de localisation des firmes étrangères et la probabilité de fondation et de fonctionnement des firmes nationales dans une ville. Les économies d'urbanisation ainsi que le potentiel de marché ont l'impact positif sur un choix de localisation. Les économies de localisation gardent la forme de U inversé.

Pour les firmes nationales de certaines industries (les industries traditionnelles et la fabrication intégrante des machines) et pour les firmes étrangères, l'impact de salaires sur la performance des firmes est positif, c'est-à-dire, la qualité du capital humain est plus importante que les coûts associés avec les travailleurs plus qualifiés. L'effet de l'infrastructure routière n'est pas robuste, néanmoins, pour certains échantillons l'effet de densité des routes automobiles est positif.

Le choix de localisation par les firmes étrangères est plus influencé par l'agglomération, le potentiel de marché et le climat économique que celui fait par les firmes nationales. Cette situation est probablement causée par les possibilités de choix plus grandes dont disposent les firmes étrangères. L'estimation montre que les firmes étrangères sont intéressées par une demande importante, autrement dit, elles poursuivrent les stratégies du marché. Elles semblent ne pas poursuivre les stratégies de rationalisation en ce qui concerne le salaire. Quant à la recherche d'actifs, cela peut avoir lieu pour une raison suivante. Si on suppose qu'il y ait plus d'innovations dans les territoires avec des activités économiques diversifiées et dans les territoires avec un climat économique favorable, l'hypothèse, que la stratégie de la recherche d'actifs est présente, ne peut pas être rejetée. Les résultats dans les Chapitres 3 – 5 confirment l'hypothèse que le climat économique au niveau régional influence la performance des firmes et leur choix de localisation pour une ville.

Ainsi, les Chapitres 3 – 5 confirment que les économies d'agglomération et le potentiel de marché sont importants pour la performance des firmes et leur choix de localisation ; les effets varient selon les types d'entreprise, les industries et les villes.

Laboratoire EQUIPPE (http://equippe.univ-lille1.fr/)
l'Université Lille 1 - Sciences et Technologies
Faculté des Sciences Economiques et Sociales
59655 Villeneuve d'Ascq Cedex

Aknowledgements

I express my deepest and sincere gratitude to Hubert Jayet, my supervisor at the University Lille 1 (France) and to Sergey Kadochnikov, my supervisor at the Ural Federal University (Russia). I am especially grateful to Hubert Jayet for his generous help, inspiring confidence and support. I am deeply thankful to Sergey Kadochnikov for brilliant opportunities and fruitful cooperation.

I would like to sincerely thank Economics Education and Research Consortium for supporting the research and providing valuable advice for the project No R09-0651 'Spatial concentration and firm performance in Russia' carried out together with Pavel Vorobyev and Nadezhda Kislyak in 2009-2010. My co-authors and I thank the panel of experts and participants of the 26th and 27th EERC workshops for fruitful comments and suggestions; we are heavily indebted to Russel Pittman, John Earle, Denis Nizalov and Volodimir Vakhitov for detailed discussion, comments, and guidance. I also gratefully aknowledge the support by Federal Agency of Education of the research project 'Inter-industry and intra-industry effects of foreign direct investment on the dynamics of spatial development: from monotown to diversified town', 2009-2011, № НК-233П (Director of the project: Igor Drapkin).

Special thanks to my colleagues Nadezhda Kislyak, Pavel Vorobyev, Pavel Kuznetsov, Igor Drapkin, Anna Fedyunina for fruitful joint work on the research projects, and to the students for their excellent assistance during the projects: Valentina Skryabina, Anna Galisheva, Viktor Grevtsev, Irina Aleynikova, Evgeny Monastyrenko, Aleksey Bolshakov, Polina Bochanova, Darya Kazakova, Maria Sazonova, Dmitry Mel'nikov.

I thank organizers of the New Economic School Summer Schools of HESP Project 'Spatial Economics: How Geography Matters For Economics' in 2010-2012, HSE-CMSSE Summer school in 'Spatial Economics and Imperfect Markets: Empirical Aspects' in 2013, and HSE-CMSSE seminars, particularly, Aleksey Savvateev, Dmitry Krutikov, Evgeny Zhelobodko, Sergey Kokovin, Dmitry Pokrovsky, Sergey Kichko, for the opportunity to listen to the brilliant lectures, have interesting discussions, and visit beautiful Russian cities. I am especially grateful to Volodimir Vakhitov for his lectures in the Summer schools, and for the valuable advice.

I am very grateful to Hubert Jayet, Sergey Kadochnikov, Jacque-François Thisse, Giuseppe de Archangelis, Fabrice Gilles, Serge Dormard, Yannick Binneau, Natalia Zubarevich, Andrey Treyvish, Pierre-Philippe Combes, Natalia Volchkova, Sergey Afontsev, Natalia Turdyeva, Kristian Behrens for very interesting lectures and valuable discussions.

I thank Wladimir Andreff, Olga Popova, Dorota Kowalczyk, Natalia Shestakova, Walid Chatti, Hakim Hammadou, Evgeny Shiltsin, Tatyana Mikhailova, Aleksandr Filatov,

Evgeny Zhelobodko, Dmitry Pokrovsky, Florian Mayneris, Frédéric Robert-Nicaud, Anna Vakhitova for valuable discussions and advice, and an anonymous referee from Journal of New Economic Association for valuable comments.

I would like to thank Michael Rochlitz for the inspiring exchange of ideas, technical advice, and interesting lectures.

I express my sincere thanks to Marion Drut, Philippe Sénave, Elena Avramenko, Sergey Lukyanov, Lyudmila Ruzhanskaya, Irina Baskakova, Elena Bedrina, Farida Gubaidullina, Elena Razletovskaya, Luo Yingchan, Léa Marchal, Kassoum Ayouba, Paul Eliot Rabesandratana, Hamza Bennani, David Desmarchelier for interesting discussions and support.

I am especially grateful to the International Centers of the University Lille 1 and the Ural State University /Ural Federal University, particularly to Véronique Level and Maria Guzikova. I would like to thank from the bottom of my heart Bernard Dupont and Véronique Level for having come to Novosibirsk, where we met five years ago. I would like to thank Carole Picault and Stéphanie Ninive for their kind help.

I am very grateful to the Embassy of France in Moscow for financing my studies in the Ph.D. program at the University Lille 1 and for the excellent administrative support during 4 years. My special and sincere thanks go to Fanny Adda, Cyliane Guinot and Pauline Perez. I sincerely thank CNOUS, Campus France, CROUS, administration of the student residence Triolo for highly valuable support in organizing my trips to Lille. I am particularly grateful to Nathalie Chawin, Mustapha Belqasmi and Saliha Abid.

I would like to sincerely thank Alliance Française in Ekaterinburg, particularly Mikhail Borisov, Olga Sokolova and Alyona Chaplygina for beautiful French lessons and for valuable assistance with translation. I thank from the bottom of my heart my family and friends, particularly Elena Eliseeva.

I express my deepest and sincere gratitude to Glenn Rayp, Oleksandr Shepotylo, Sergey Afontsev and Lise Patureau for their valuable comments and ideas for further research.

Contents

List of figures

List of tables

Introduction

In Russia, as in the other countries, regions differ in initial conditions and in economic development; differences among cities within regions are substantial as well. However, the science of economics is concerned with providing high standard of living for all people in the country, having the limited resources. In the situation of competition, opportunities for development differ among regions and cities. Policy alternatives in their extremes are either to support geographic concentration of economic activities in clusters or large cities or to promote equal geographical distribution of economic resources. To justify the prevalence of one or another of these directions, it is necessary to be aware of the amount of possible economic externalities, and to keep in mind social consequences.

In the Soviet Union urbanization was artificially limited; population was 'attached' to the periphery. Now mobility remains low in Russia. However, transition towards market economy gave people an opportunity to 'vote with their feet', favouring centripetal tendencies in the country. Among the factors necessary for successful development of territories are population density, real household disposable income and well-balanced industrial composition. For a country in general and for the territories within the country, the background for development are climate, market capacity, and communications and transport infrastructure. In this research we are interested in the factors of development on the city level. Therefore we take into account size and diversity of the cities, concentration of particular industries, and home market potential of cities. Besides, regional infrastructure is considered, i.e. road infrastructure and regional business environment risks. This research provides some insight into the impact of agglomeration on enterprise performance and on enterprise location choice for a city. Special attention is paid to the sources of economic growth and development, specifically, the nature of agglomeration economies, arising on the supply and demand side.

Diversity, openness, and exchange of ideas are among the factors determining economic growth and development. Personal development is closely linked to the productivity growth, and therefore to economic development of a country. It is well-understood that the perspectives of exploiting natural resources potential are quite limited, and modernization of industries and institutions is needed. There is connection between modernization of the economy and diversity of cities. According to the life cycle theory (Vernon, 1960), new industries benefit from location in diversified places. An important aspect is the impact of diversity on productivity. The recent crisis showed that mono-specialization of regions on resource-oriented and exclusively export-oriented industries

makes them vulnerable. Therefore economic diversity is becoming an essential element of regional strategic planning; particularly, providing legal and financial preferences and other guarantees for the investors (Expert, 2008). In this study, diversity within manufacturing sectors is considered, particularly, the dynamics of diversity in the Russian cities over the recent years and the impact of diversity on business performance and location choice for a city. Following further the life cycle theory, mature industries benefit from localization economies, i.e. they are better off in case of location in highly specialized cities. In this work we estimate the economies firms get from location in the cities where a lot of similar firms already work. We also touch upon the advantages and problems faced by the highly specialized cities, the 'monotowns'.

In this work the idea of diversity implies the variety of business activity in cities, and is extended to consider the impact of foreign direct investment on business performance and location choice. People from different backgrounds, while working together, acquire a unique opportunity to exchange tacit knowledge; ideas flow and the new technologies arise. Variety of cultures stimulates capability of mutual understanding, this way forming a social and economic climate favourable both for individual well-being and for economic development. Flexibility, adaptability, diversity are the features associated with opportunities. Foreign direct investment (FDI) stimulates economic development and new job creation in the host countries and more efficient production chains. However, both large cities where diversity flourishes, and FDI that bring in new technologies, have their drawbacks. Congestion costs are associated with large cities; unemployment in the FDI home-countries arising from off-shoring of production, and controversial effects of competition with more advanced foreign firms for the local firms in the host country are among them. Therefore, these issues deserve closer insight.

The following issues are considered in this research. In Chapters 3 and 4 the impact of enterprise agglomeration and home market potential on the enterprise productivity in a city is studied, in Chapter 4 agglomeration effects are subdivided into those arising from concentration of national and foreign firms. Particularly, the impact of foreign firms' concentration on the national firms' performance is considered. To account for various types of agglomeration economies, we study diversity, urbanization and localization economies. Human capital and infrastructure are taken into account. The hypothesis is that both agglomeration and home market potential positively affect enterprise performance. Localization economies are expected to have an inverted U shape. Another hypothesis is that presence of foreign firms is favourable for the performance of enterprises in a city, as they bring in new approaches to work, and are generally more efficient than national firms.

A question arises, how to measure agglomeration? In this work, a number of indices are analyzed, and a variation of diversity index proposed by Vorobyev et al. (2010) is used. Applying this and the other indices, localization, urbanization and diversity in the Russian cities are measured. The effects of agglomeration on the enterprise performance are estimated. Another question of interest is the factors determining location choice of enterprises for a city. In Chapter 5 the issue of the territories' investment attractiveness is studied. It is another question if firms really choose where to locate, or there is a certain 'path-dependence' of business allocation in Russia, which in the times of planned economy was located based on the reasons other than economic efficiency. As foreign firms seem to have more choice in location, factors making a city attractive for the national firms and for the foreign firms were analyzed separately. The 'probability of location choice' studied in this chapter can be interpreted for the foreign firms as actually a probability of location choice for a city, and for the national firms as probability of formation and work in a city. Factors determining location choice of enterprises with different productivity levels and different age are analyzed.

Concerning methodology, to study the impact of economic concentration on enterprise performance, panel data model was estimated, using the firm level data for 1999-2008. In Chapter 3, where the impact of spatial concentration on the firm performance was analyzed, fixed effects were used to deal with endogeneity arising from unobserved heterogeneity among firms. Endogeneity arising from simultaneity was assumed not to be present for the reasons explained in Chapters 2 and 3. In Chapter 4, where spatial concentration was subdivided into concentration of national and foreign firms, first differencing was used to deal with endogeneity arising from unobserved heterogeneity. Instrumental variables (agglomeration variables with 2 period lags) and GMM were used to deal with endogeneity arising from simultaneity. In Chapter 5 conditional logit model was applied; data for the year 2007 was used.

The existing research on the topics studied here will be considered in the next chapter. Briefly, nowadays spatial economics attracts interest as New Economic Geography, starting with the works by Krugman (1991), Fujita (1988) and Venables (1996). Marshall (1920) assumed that among the reasons behind industrial concentration were 'the mysteries of trade'. Agglomeration economies in the city are associated with concentration of similar firms (localization or specialization) or of firms belonging to various industries (diversity interrelated with urbanization). Agglomeration economies are based on 3 principal mechanisms (microfoundations): sharing, matching and learning (Duranton and Puga, 2004). For instance, 'learning mechanisms' are mainly a source of diversity economies. In Russia internal scale economies dominate, i.e. enterprise productivity is affected by its size (Gonchar,

2010; Russian Economic Report, 2007). For transition countries, there is some evidence of positive agglomeration externalities, but overall, the results of the existing research are mixed (Beaudry and Schifaerova, 2009).

The main results. In Chapter 3, urbanization and diversity economies were found to be significant and positive. Localization economies were found to have an inverted U shape for the majority of specifications, but to be positive for monotowns, showing that concentration of the firms in the same industry is particularly important for the enterprises' performance there. Urbanization economies proved to be particularly important for monotowns too. Home market potential is positive and significant. In Chapter 4, it was found that urbanization economies arising both from national and foreign firms are positive, while those arising from foreign firms are relatively larger in the model with first differencing and GMM. Localization economies arising from national firms have inverted U shape, while localization effects of foreign firms' presence are contradictory. In Chapter 5, enterprise location choice for a city was estimated. Both urbanization economies and home market potential are significant and positively affect location choice; localization economies keep the inverted U shape. Foreign firms' choice for a city is the more affected by agglomeration level, home market potential and institutional infrastructure, probably because they have greater opportunities to choose location.

This work is organized as follows. The first chapter reviews the theoretical and empirical literature on agglomeration economies. Chapter 2 outlines the methodology applied in this study, specification of agglomeration indices and classification of industries and territories. In Chapters 3 and 4 the impact of agglomeration and home market potential on enterprise performance is studied. Chapter 4 goes into further details of these factors, studying the impact of the national and foreign firms' concentration on the enterprise performance. In Chapter 5 factors determining location choice of enterprises for a city are analyzed. The conclusion follows.

1. Literature review

Literature review is structured the following way. In Section 1.1, literature concerning the impact of agglomeration and home market potential on enterprise performance is reviewed. Section 1.2 is devoted to literature on foreign direct investment effects. In each section, firstly, theoretical background is reviewed, followed by the main conclusions of empirical works and research on transition countries, including Russia. Section 1.3 is devoted to business environment and regional policy in Russia.

1.1. The impact of spatial concentration on enterprise performance

Theoretical background. Spatial economics has represented a powerful branch of economic theory, it dates back to Weber (1929), and nowadays it attracts interest as New Economic Geography, starting with the works by Krugman (1991), Fujita (1988) and Venables (1996). The research focus has been changing to reflect new economic realities. In the past researchers were occupied mainly with the optimal spatial allocation of resources that would minimize transportation costs. Later interest moved to the phenomenon of agglomerations associated with large cities or industrial clusters attracting human, natural, and financial resources. New economic geography explains agglomeration combining trade costs with scale economies (Krugman, 1991).

Agglomeration in a broad sense can be defined as concentration of economic activity in certain territories (Marshall, 1920). Krugman (1991) draws our attention to spatial organization of economic activity in the United States, where the majority of population resides in the East Coast, although climate there is not the most favorable, and to the pattern of economic activity location in Europe, where nighttime satellite photos indicate concentration of activity in or near Belgium. Marshall (1920) analyzed the supply side of agglomeration economies. He assumed that among the reasons for concentration of economic activity were 'the mysteries of trade' that became accessible to people in places of industrial concentration. Following Neffke (2009), we define agglomeration externalities as positive and negative effects of scale and scope resulting from economic activity concentration. Agglomeration economies play an important role in geographic allocation of resources because firms are likely to take them into account while choosing their location.

New economic geography (NEG) models rely on the concept of home market potential. If two-region economy and the industries with imperfect competition are considered, then more than proportionate share of enterprises choose to locate in a region where local demand is larger (Ottaviano, Thisse, 2004). Location choice of enterprises is determined by various factors. There are factors inherent to place that determine its attractiveness for enterprises, such as geographical location and natural resources that can be classified as factors of the first nature. Besides, there are factors created by people, long term capital investments such as those into transport infrastructure etc., i.e. factor of the second nature. Finally, there are agglomeration externalities arising from concentration of economic activity in certain place, i.e. factors of the third nature (Mikhailova, 2013).

Rosenthal and Strange (2004) determine 3 scopes of agglomeration economies: industrial, geographical and temporal ones. In this research we focus on the industrial scope. It reflects the industry boundaries of agglomeration economies: whether agglomeration

economies are localized in one industry or they arise when different industries interact. In this respect, two types of agglomeration economies are known in the literature. Localization economies (MAR-externalities, firstly developed by Marshall (1920) and rediscovered by Arrow and Romer) are industry-specific externalities. They are associated with specialization – high concentration of economic activity in the same industry in an area. Diversity economies (Jacobs' externalities) are industry-universal externalities. They imply cross-fertilization of different industries in an area and were explicitly formulated by Jane Jacobs (1969). They are externalities from diversity of economic activity in an area outside the own industry. Urbanization economies are associated with the city size – the number of people or the volume of economic activity in a city (Rosenthal and Strange, 2004). Diversity and urbanization economies are closely linked.

Diversity (urbanization) effects related to production and consumption are scale economies, common input market, decrease of transaction costs, additional opportunities, such as lower probability of unemployment and highly diversified goods and services (Quigley, 1998). Jacobs (1961) emphasized the importance of industrial diversity for social security and urban economic growth. She claimed that small firms benefit more than the large ones from urban diversity, due to their stronger dependence on the market situation. In the other work Jacobs (1969) demonstrated that urban economic diversity is a critical condition for innovation.

Concerning economic foundations behind agglomeration processes, agglomeration economies are based on 3 principal mechanisms, or microfoundations: sharing, matching and learning (Duranton and Puga, 2004). Sharing mechanisms imply that agglomerations facilitate sharing of some common indivisible resources by firms, including infrastructure, variety of intermediate inputs, specialized labor, and also risks. Matching mechanisms account for creation of pools of specialized workers; so the costs associated with training of workers decrease. Finally, learning mechanisms (technological and knowledge externalities) refer to the possibility of more intensive innovations in diverse environment of agglomerations stimulated by information spillovers (Duranton, Puga, 2004). Localization economies are explained by all three microfoundations. Diversity economies are based primarily on learning mechanism. They are associated with demand as well as supply and explain advantages of life in a large city with its variety of goods and services (Ottaviano and Thisse, 2004). Sharing externalities are in action both for localization and diversity measures. Matching mechanisms of agglomeration economies are reflected by the majority of localization indices, and to a less extent by diversity indices. Diversity indices are more suitable than the localization ones to account for learning mechanisms.

Audretsch (1998) concludes that location of firms in a close proximity remains important in spite of development of telecommunication technologies, partly due to tacit knowledge. Dissemination of such type of knowledge takes place during face to face communication, which is more intensive within agglomerations (Elison and Glaeser, 1997). These effects are particularly evident in the university cities, for example, in USA, where they became centers of industrial development for such branches as semiconductors, biotechnologies, software. Tacit knowledge is important in the contemporary times of communication technologies, as these technologies are complementary with face to face communication (Robert-Nicaud, 2013). Another aspect of agglomeration effects are that people in large cities are more specialized, and therefore more productive, according to the logic of economic specialization suggested by Adam Smith (intensive margin). Besides, people are more numerous in the large cities, therefore they perform more tasks, becoming better as a team (extensive margin) (Robert-Nicaud, 2013).

Theoretical and empirical research on urban development concludes that there is a positive correlation between city size and wages, after controlling for ability. Explanation behind this is threefold. Firstly, there are high living costs in the cities, so wages are compensation. Secondly, highly qualified workforce is attracted to the cities, as a result of sorting. Thirdly, there are agglomeration externalities generated by interaction between the qualified employees and between the firms. These are the agglomeration economies mentioned above: input sharing, matching and learning. As a result, productivity grows. Under imperfect competition wages can be defined as productivity minus profit margin of firms, and (assuming that workers bear training costs to acquire the type of qualification required by the firm), minus training costs. So, wages are proportional to productivity (output per capita), and as productivity grows with city size, wages grow as well (Robert-Nicaud, 2013). In the research presented in the following chapters, wages are included as an explanatory variable for enterprise performance and location choice, though they are expected to be closely related with agglomeration indices.

Along with positive agglomeration effects there are drawbacks associated with high concentration of economic activities, namely, with large cities, such as crowding, high transport costs, ecological and social problems. There is a hypothesis of an inverted U-shape of agglomeration economies, proposed by Mills (1967) and Mirrlees (1972). Mills (1967) develops a general equilibrium model with imperfect competition (monopoly in the goods market) and increasing returns to scale. He points out a possibility of congestion in the city beginning with a certain population density due to transportation costs and costs of housing. It implies diminishing agglomeration economies after having reached a certain level of

concentration, in other words, it reflects the hypothesis of inverted U-shaped agglomeration economies. In this work, the inverted U shape of localization economies is assumed.

Analysis of business agglomeration is interrelated with cluster theory. Porter points out competition as one of the development factors within cluster. According to him, FDI is another element important for cluster development (Porter, 2000). Vorobyev et al. (2008) classified industrial clusters in various countries into the models, based on their institutional features. Each model represents combination of 6 key characteristics of cluster: the degree of market connections and competition, presence of firms-leaders, development of small business, innovations, internationalization, and presence of FDI. FDI plays the key role in the Indian and China model, bringing in modern technologies, investment and facilitating entry onto the world markets. In the Soviet-Russian model market relations and competition are minimal, and production is concentrated in the large firms. As for the Italian model, industrial cluster consists of a large number of small firms united into various associations for enhancement of competitiveness.

Review of empirical results. Results concerning agglomeration effects are contradictory. In the recent overview of empirical works Beaudry and Schiffauerova (2009) document that researchers have found only positive MAR-externalities in 51 cases, in 11 cases – both positive and negative ones, and in 20 cases they were non-significant. Only positive Jacobs' externalities were found in 56 cases and non-significant in 46 cases; both positive and negative effects were found in 13 cases, and only negative effects in 9 cases. De Groot et al. (2009) analyzed 31 papers and also found divergence in empirical results: both types of agglomeration externalities are positive in about the same number of cases as they are negative. For instance, they found that about half of the authors reveal negative Jacobs' externalities. De Groot et al. (2009) have performed meta-analysis to study the determinants of the signs and significance levels of parameter estimates and came to the conclusion that the choice of the dependent variable, the control variables, and the construction of agglomeration indices affect their signs and significance levels. Differences between countries and time periods also affect the parameter estimates. All this makes estimation of agglomeration effects context dependent (Neffke, 2009). In our research we distinguish between the groups of industries, types of cities, types of enterprises, and introduce enterprise fixed effects to enhance accuracy of the empirical model; however, the issue of context dependence cannot be avoided completely.

There are the following examples of evidence of positive agglomeration effects. Productivity premium received due to doubling of a city size, expressed through the population of a Standard Metropolitan Statistical Area (SMSA) for USA varies from 3 to 8%

(survey by Rosenthal and Strange (2004)). According to Sveikauskas (1975), for an average industry, doubling of city size leads to productivity increase of 5.8%. In the study by Okubo and Tomiura (2010), productivity premium from being in a core region versus periphery was found to be 20-50%. Based on the US data for the year 1988, Ciccone and Hall (1996) found that doubling of employment density in a county leads to 6% increase of average labor productivity, where density is measured as the amount of labor per square foot; the results are aggregated from the county to state level. Based on data for Great Britain, the years 1998-2003, Anastassova (2006) found that estimated elasticity of hourly earnings with respect to employment density (the average number of full-time employees per square kilometer in a given district) is around 4 percent, both for district and county levels. The role of agglomeration economies and primarily knowledge transfers is emphasized by the evidence for the US metropolitan areas, years 2001-2005; productivity rises by 2 or 4 percent if density of human capital doubles, while human capital positively influences this process. Density of human capital reflects interrelation of skill and density, the latter accounting for the spatial distribution of population within metropolitan areas, i.e., the average number of full-time employees per square kilometer (Abel et al., 2010).

Applying plant-location fixed effects along with industry-time fixed effects, Henderson (2003) finds that localization externalities exist in high-tech, but not in machinery industries, and that diversity economies exist in machinery industry – corporate sector. Martin et al. (2011) based on French data on enterprises for the years 1996-2004 find benefits from localization economies, but no benefits from diversity economies or competition effects. They find the returns of localization economies to total factor productivity (TFP) to be 5-10%. They also find that firms internalize benefits from clustering, and therefore important effects from cluster policy should not be expected in developed countries. The dependent variable in their research is value added; explanatory variables are production factors, localization, diversity, urbanization and competition coefficients (in logarithms); they use time differencing approach. Time, employment area, industry and firm fixed effects are applied. While Henderson (2003) finds that it is the number of firms but not employees that generates externalities, Martin et al. (2011) come to the opposite conclusion.

Henderson et al. (1995) find evidence of Marshall-Arrow-Romer (MAR) externalities in mature industries and evidence of both MAR and Jacobs' externalities in newer industries. These conclusions are in line with concepts of agglomeration and product cycles. At the initial stage of development an industry benefits from a diverse environment within a metropolitan area. At a later stage, when the industry matures it benefits from location in smaller specialized cities. Localization economies dominate in such industries as textile,

apparel, transport equipment, primary metals, food processing, pulp and paper. Diversity (urbanization) economies dominate in such industries as high-fashion apparel, upper-end publishing, and business services; they benefit from location in very large metropolitan areas.

At the enterprise level, diversity effects are usually found to be significant for new enterprises, while localization externalities are important for mature enterprises with standardized manufacturing production (Duranton and Puga, 2001; Neffke et al., 2008). Overall, 'the nursery cities' model (Duranton and Puga, 2001) implies that diversified cities and specialized cities co-exist, but are associated with different stages of industrial development. Within the product cycle concept, it is assumed that firms move from an exploration (or search) stage to exploitation (or mass production) stage; at the first stage of development location in the diversified environment is favorable for firms, while at the second stage firms are more efficient when located in the specialized cities (Neffke, 2009).

There is other evidence that localized sectors are mainly "traditional" sectors (like jewelry, wine, and textiles) and sectors where scale economies are important; however, once countries' industrial structures are controlled for to take into account numerical prevalence of traditional sectors, science-based sectors turn out to be the most localized ones (Vitali et al., 2009). It leads to a conclusion that a possible explanation for the mixed results is the industrial characteristics, including the concept of product cycle. It is important to analyze specific industries in order to make conclusions for the cluster policy, because it usually deals with a certain industry and therefore requires information exactly about agglomeration economies for it.

Evidence for transition economies. The Institute for Analysis of Enterprises and Markets, Higher School of Economics (IAEM) in cooperation with World Bank, IMF and 'Levada-Center' has carried out research on Russian enterprises based on a survey of 1000 enterprises of eight manufacturing industries in 156 cities of 50 Russian regions. For the majority of manufacturing industries in Russia efficiency of enterprises, located in the cities with population less than 50,000 people, was found to be lower than that in the large cities. The value of output produced by an employee of a manufacturing enterprise in mega-polis is 1.2 mn roubles, in large and medium-sized cities – 900,000 – 1 mn roubles, and in small cities – about 800,000 roubles per year. According to this research, doubling the city size leads to 5% increase in productivity (Enterprises and markets, 2010).

Overall, manufacturing in small cities is less competitive than elsewhere. It happens due to a number of reasons. Apart from out-of-date technologies and being far away from final goods markets, plants and factories often face another problem - social responsibility - that makes firing unneeded personnel impossible. They bear the burden of 'non-profile assets'

– community services, transport, kindergartens. However, the situation varies greatly across industries. For enterprises of textile and electro-optical industries located in small cities productivity is 1.5 times less than average. However, wood-processing enterprises in small cities have productivity 2% lower than in large and medium-sized cities, and productivity of enterprises in chemical industry is even 25% higher than average. Productivity of certain machine-building enterprises in small cities is also high. According to the experts, enterprises in small towns and monotowns can be competitive if there are a lot of other enterprises, including those of small business, located in the same territory (Enterprises and markets, 2010).

Gonchar (2010) explores the impact of city size and urban agglomerations on company productivity and growth in Russia based on the survey of 1,000 manufacturing enterprises of 168 four-digit industries (eight two-digit) in 49 Russian regions, conducted in 2005-2006. She finds that enterprises belonging to agglomerations are on average 46% more productive than the others. Enterprises in cities with less than 50,000 inhabitants close to an agglomeration center are 37 percent more productive than the similar ones in isolated cities. Enterprises in cities with 50,000-250,000 inhabitants are 61 percent more productive. Concerning industrial differences, wood-processing, food and transport machinery industries benefit the most from agglomeration effects; there are no agglomeration effects in chemicals and steel industries.

Gonchar (2010) estimates a probit model, where the dependent variable is company productivity; explanatory variables include the city, regional, and enterprise characteristics that might affect enterprise productivity. Regressions are estimated for the probability to have productivity above the industry average (2-digit sector) and probability that enterprise would belong to the group of competitive firms. She finds that internal scale economies dominate in Russia, as the size of an enterprise affects productivity. She also finds external scale effects: the urban agglomeration positively affects value added productivity. Besides, Gonchar concludes that infrastructure and transport system development would allow smaller cities to gain from agglomeration effects by improving communication between agglomeration center and smaller cities. Along with agglomeration effects, regional integration into the world economy proved to be important.

Concerning specific features of the Russian agglomerations, firstly, the majority of the urban population lives in the agglomerations, but from the point of view of rational organization of the economic space, the number of cities forming the agglomerations is too low. Secondly, a lot of large cities do not generate agglomeration effects due to low diversity of functions and resources, and because of underdeveloped transport infrastructure, which

would be supposed to connect satellite cities with the center. Thirdly, a lot of industrial agglomerations are at the interim stage of development. Mature agglomerations in Russia are located mainly in the central-European part of the country, in the Volga region, in the Urals, and around a few cities along the Trans-Siberian railway (Gonchar, 2008). Moreover, Lappo and Polyan (2007) showed that in the USSR a lot of cities were created based mainly on the industrial needs. However, many of these cities could not form well-functioning agglomerations, because diversity within them is not enough to create agglomeration economies. In a number of cities (Ivanovo, Chelyabinsk, Volgograd, Bryansk, Tyumen, Lipetsk, Kurgan) production remains highly specialized (Lappo, Polyan, 2007).

Among the works on the Former Soviet Union and other transition economies that distinguish localization and diversity (urbanization) externalities are the works by Vakhitov (2008) and by Vakhitov and Bollinger (2010). They use establishment level longitudinal data on Ukrainian machine manufacturing and high tech industries for the period 2001-2005. Research is based on the empirical model by Henderson (2003). The authors apply a standard log-log production function for a given establishment *i* in the area *m*, industry *s*, at time *t*, where dependent variable is logarithm of the firm's output, and the independent variables reflect resources used by firm, agglomeration indices, institutional variables, and industry-, time-, and QMSA ('quasi-metropolitan' statistical areas) fixed effects.

Vakhitov (2008) and Vakhitov and Bollinger (2010) conclude that localization and diversity (urbanization) externalities exist both in machine manufacturing and in high tech industries; the magnitude of both types of externalities is larger for high tech sector. According to Vakhitov and Bollinger (2010) it is privately-owned firms, but not state-owned firms, that are able to benefit from agglomeration effects; agglomeration economies of foreign-owned firms are the largest ones. The authors conclude that the capability to gain from agglomeration economies depends on management level. Vakhitov and Bollinger (2010) claim that although in the Soviet Union enterprises were formed according to central planning, growth of cities around the plants, and vertically integrated 'territorial-production complexes' generated some localization and diversity economies. Vakhitov (2008) emphasizes that agglomeration effects are stronger on the local level and diminish with distance.

Research by Cheviakhova and Rytchkov (2004) is among very few works comparing spatial concentration of economic activity under centrally-planned and free market systems. They apply the framework of New Economic Geography using regional structure of industrial employment and find systematic and persistent changes in regional pattern of industrial employment in Russia during 1985-1995, particularly, in the beginning of 1990ies; they also

find that these changes are explained in the framework of New Economic Geography. The authors construct New Economic Geography Factor (NEGF) - the difference between the initial point and the equilibrium, expressed using the vector of the number of industrial firms and wages in the region and reflecting a combination of the model inputs - the distance between regions, population of regions, initial industrial shares and wages. NEGF accounts for the direction towards the equilibrium and predicts about 15% of total variation in the shares of regional industrial employment between the years 1985 and 1995. Therefore the authors conclude that restructuring took place due to market forces.

Golovanova (2008) analyzed changes in spatial structure of manufacturing industries in Russia in 1998-2004 based on Rosstat data. She applied Herfindahl-Hirschman index and revealed that in 23 out of 202 industries (OKONKH classification) spatial concentration increased, and in 10 industries it decreased; in 20 industries spatial production structure changed. She concluded that in Russia the tendency for production centers formation is stronger than the tendency for decentralization; spatial concentration is more pronounced for the industries, facing stronger competition from foreign producers and weaker for exporting industries; analysis of wood-processing industry and manufacture of pulp, paper and paperboard revealed increasing spatial concentration of technologically related industries.

There are a few works estimating the factors of growth on the regional level. According to the World Bank Report (Russian Economic Report, 2007) for Russia, two types of agglomeration effects are the most important: the size and scope of cities (urban agglomeration) and spillovers on growth from one region to another (spatial correlation of growth). There are spillovers from proximity to rapidly growing regions with large economies, particularly around Moscow, St. Petersburg and Rostov in the South (Russian Economic Report, 2007).

The above mentioned studies on Russia operate the indicator of city size which embeds both localization and diversity dimensions. The question of distinguishing between localization and diversity externalities is among the central ones in our study. While the research of IAEM sheds light on the industrial differences in agglomeration economies, we emphasize their importance as well - in the theoretical part of our study we describe specific features of production causing these differences.

In the research by Vorobyev et al. (2010) localization and diversity economies in Russian cities were analyzed based on data for more than 3000 Russian manufacturing firms. Performance was measured by the growth rate of enterprise revenue during 2001-2004. Differences between 5 classes of industries, and between new and old generations of enterprises were analyzed. Significant inverted U-shaped localization economies and positive

economies from diversity of industries in a city were found on 3-digit level of industrial classification (OKVED), 2-digit level and using classification of industries into 5 groups. With 2-digit level of industrial classification, inverted U-shaped localization economies were found for basic materials and science-based industries, and linear positive economies were found for traditional industries. It was also concluded that localization economies are not fully utilized. A diversity index, variation of which is used in this research was introduced by Vorobyev et al. (2010).

Home market potential is analyzed based on the approach by Combes et al. (2008a). Concerning the demand side of location decisions, Krugman (1980, 1991) shed light on the role of market size in attracting enterprises to the region. He constructs a Core-periphery model and shows that enterprises are driven to the large regions by the demand existing there, which allows saving transport costs, i.e. enterprises are attracted by the Home market potential. Baldwin and Okubo (2005) introduce Melitz (2003) approach with heterogeneous firms into New Economic Geography model. They reveal that firm heterogeneity diminishes Home market potential and that it is more efficient firms that are attracted by large regions. Firstly, it implies 'self-selection' of more effective firms into the core region and overestimation of agglomeration externalities in empirical research. Secondly, it implies 'sorting effect' of production subsidies, i.e. attraction of the least productive firms to the periphery by regional policy measures.

Overall, there is some evidence of positive agglomeration externalities in transition countries, although the results are contradictory. Our contribution to the existing research is to explain enterprise performance and location choice based on agglomeration and home market potential, using the firm level data and city level analysis. Following the research described in this section, we subdivide agglomeration into localization, urbanization and diversity; we take into account infrastructure of the regions where cities are located.

1.2. The impact of FDI concentration on enterprise performance

Theoretical background. Importance of foreign direct investment (FDI) as a form of foreign market entry for the firms has been growing over the recent years. Empirical research shows that multinational corporations (MNCs) that own or control subsidiaries, or real and financial assets in at least two countries (Andreff, 2003) i.e. are involved in FDI tend to be more productive than the other firms. They possess advanced production and managerial technologies and employ qualified personnel. It is in line with the Ownership advantage within the concept by Jonh Dunning on Ownership, Location and Internalization advantages

of MNCs. It makes them able to bear high entry costs associated with foreign markets and to compete with local firms that have in their turn more information and connections on the local market (Mikhailova, 2009).

Presence of firms with higher productivity can positively affect domestic firms, especially in the close proximity (Bode et al., 2009, Drapkin et al., 2012). Development agencies and policy makers tend to assume that spillovers from FDI will positively affect economic development of the country, which makes attraction of FDI an important policy issue[1] (Smarzynska-Javorcik, 2004). There is indeed evidence that FDI is associated with external effects through productivity spillovers to domestic firms, presumably generating factors of growth for the host economies[2]. However, there is evidence both on positive and negative FDI effects, making the impact of FDI on the host country contradictory (Smarzynska-Javorcik (2004), Stancik (2007), Merlevede and Schoors (2008)).

FDI can be defined as investment with the aim to organize a subsidiary in a foreign country or to take control over an existing firm abroad. A key feature of FDI is long term economic relations with a business unit established abroad. A minimum share in a foreign company for investment being the 'direct' one is defined relatively arbitrary. According to the United Nations, World Trade Organization, International Monetary Fund, OECD, INSEE, Bank of France, it is 10% of capital in the foreign subsidiary (Andreff, 2003). According to the Russian legislation, namely, the Federal law 'On the foreign investment in the Russian Federation', this share is also 10%[3]. FDI implies that the 10% share consists of the voting stock (ordinary shares), providing investor with opportunity to control the company. FDI can be carried out by organizing a subsidiary, by acquiring shares of a foreign company, by setting up a joint venture, by mergers. FDI exists as greenfield projects, i.e. creating new

[1] Smarzynska-Javorcik (2004) cites the following cases: during the late 1980ies, in the state of Kentucky Toyota received an incentive package, which in present value was worth 125-147 million dollars, the expected employment of the plant being 3000 workers. As for transition countries, during 2000 in Hungary 92.6% of all tax concessions was provided to foreign investors. Stancik (2007) points out that in Czech Republic the system of subsidies for foreign investors was approved in 1998.

[2] Markusen and Venables (1999) cite an example of economic development in South-Eastern Asian countries – Taiwan, Hong-Kong, Singapore. Initially investments were carried out by TNCs into the industries not developed in the country: electronics, household appliances, consumer goods, and there were no local firms in these industries. Entry of TNCs into the market provoked a *positive backward effect* that led to development of the industries – suppliers of resources and services for TNCs. Afterwards, efficient local companies appeared in the industries created by the TNCs, and there is evidence that in some cases TNCs were crowded out. Similar kinds of examples exist for China, where comparative advantage moved from traditional sectors (such as production of clothes and carpets) to the high tech sectors (computers, electrical appliances, and electronics).

[3] Federal law 'On the foreign investment in the Russian Federation' N 160-FZ from 9 July 1999. Database of legal documents 'Consultant' http://base.consultant.ru/cons/cgi/online.cgi?req=doc;base=LAW;n=121824;fld=134;dst=4294967295;rnd=0.020425477530807257;from=16283-0

operational facilities in a foreign country, or as foreign capital inflow to an existing domestic firm (Stancik, 2007).

FDI effects occur as direct effects and productivity spillovers (external effects). Direct effects are the changes within the firms – FDI recipients (Griffith and Simpson, 2003; Damijan et al, 2003). Along with financing, FDI brings in new equipment, new technologies – product, process, and managerial ones (classification in Kadochnikov and Esin (2006)). Firms can gain skills related to new technologies – via training of labor force and transfer of tacit knowledge within the firm. The advantages that the firms receive with FDI make them more competitive; possible increase in production allows benefiting from scale effect. As a result, direct effects are assumed to stimulate production growth.

As for the domestic companies that do not receive FDI, they are affected by FDI via productivity spillovers, such as competition effects, technology transfer effects, and improvement of workforce qualification (Kadochnikov and Kulakova, 2004). Markusen and Venables (1999) point out that foreign direct investment (FDI) may affect a host economy through product market competition and linkage effects. Productivity spillovers occur if the foreign firms' entry leads to productivity increase for domestic firms in the host country, and the value of these benefits is not fully internalized by the foreign firms (Smarzynska-Javorcik, 2004). Possibility of technological transfer is among the reasons why governments of countries – FDI recipients aim at attracting FDI (Blomström and Kokko, 2003).

Spillovers from FDI can be subdivided into horizontal, vertical and also supply-backward ones. Horizontal spillovers are the effects that take place in the industry receiving FDI. They include demonstration effect – diffusion of technologies via exhibitions, educational programmes etc.; technological improvements – purchased technologies, innovations by the firm itself, technologies adopted from foreign firms; competition effect that creates incentives to increase effectiveness, and crowding out effect. However, foreign firms are certainly interested in protecting their know-how. They do it via formal protection of intellectual property, trade secrecy, paying higher wages, or choosing the countries or sectors where capacities of domestic firms to copy their technologies are low (Smarzynska-Javorcik, 2004).

Vertical spillovers are inter-industry effects that arise between firms - buyers and suppliers (Smarzynska-Javorcik, 2004). Vertical effects are backward spillover effects (effects on local suppliers of foreign affiliates) and forward spillover effects (effects for buyers of the intermediate product that was produced by firms with FDI). As for backward spillovers they arise through direct knowledge transfer from foreign firm to its domestic supplier; higher requirements stimulating quality and timely delivery; indirect knowledge transfer via copying

the technologies of foreign firms by observations or hiring workers previously employed by foreign firms; increased demand for intermediate products allowing suppliers to benefit from scale economies; competition effects that can occur among the domestic suppliers as foreign firms can choose to buy from suppliers abroad (Smarzynska-Javorcik, 2004).

Local suppliers have incentives to improve technologies in order to meet the requirements of firms with FDI – consumers of intermediate goods and services for business. Foreign companies are also interested in assisting local suppliers to improve their productivity. Smarzynska-Javorcik (2004) emphasizes that in case of backward effects arising from technology spillovers it would be reasonable to subsidize FDI, but in case of backward effects based on competition, it is not, as there are other ways to increase competition, such as trade liberalization and antitrust measures. Another effect similar to backward spillovers is development of financial institutions and business services interrelated with presence of foreign firms in the country. There were also found effects arising from firms with FDI over domestic suppliers to domestic firms using domestic inputs, supply-backward linkages (Merlevede, Schoors, 2008).

Overall, MNCs are interested to protect information and technologies from their competitors in the same industry in the host country. The same time, they are interested in high quality intermediate goods, i.e. in development of industries within their production chains. Therefore vertical FDI spillovers (backward linkages) are more likely to be positive than the horizontal ones (Smarzynska-Javorcik, 2004).

In the framework of agglomeration economies, horizontal spillovers are related to localization effects as they account for the situation when a firm or firms affect businesses belonging to the same industry (Blomström and Kokko, 2003; Markusen and Venables, 1999; Markusen and Trofimenko, 2007). Concerning vertical spillovers, concentration of firms with FDI among firms – buyers or suppliers has an impact along the supply chain. Besides, presence of firms with FDI stimulates development of the existing firms and creation of new firms in the other industries, including business and financial services. Then variety of intermediate and final goods increases, i.e. vertical effects lead to higher economic diversity (Markusen and Venables, 1999; Altomonte and Resmini, 2001). Moreover, diversity effect can originate from increased demand stimulated by growth of firms' average productivity, and higher wages resulting from it. Overall, vertical spillovers are partly interrelated with diversity and urbanization effects. Competition effects associated with agglomeration of economic activity can also be considered in relation with FDI (Markusen and Venables, 1999).

In this research, we find positive effects arising from presence of the foreign firms in the industries other than that where a firm works (urbanization effects). The effects from presence of foreign firms in the same industry (localization), associated with horizontal spillovers, turned out to be contradictory, and in most cases negative.

FDI effects can be considered in the framework of growth theories. Balasubramanyam et al. (1996) subdivides growth theories into post-Keynesian growth models, where the role of savings and investment in promoting growth is emphasized (Harrod-Domar type models), neo-classical models emphasizing technical progress (Solow type models), and new (endogenous) growth models emphasizing the role of research and development, human capital accumulation and externalities (Romer-Lucas type models).

As models with heterogeneous firms are used in the analysis of FDI, below a brief overview of these models is provided. Models that take into account firms' heterogeneity, lead to deeper understanding of the changing world trade patterns, and also of FDI and production reallocation (Melitz, 2003; Bernard et al, 2003; Bernard et al, 2003a; Yeaple, 2002; Melitz and Ottaviano, 2008; Helpman et al., 2004; Helpman et al., 2008). Analysis of firms' heterogeneity, mainly in terms of productivity, extends the framework of classical, neoclassical and modern theories. It made possible study of determinants of the firms' involvement in foreign economic activity (positive analysis) and new aspects of the economic openness impact on economic welfare (normative analysis) (Drapkin et al., 2012).

The following facts are taken into account in models with firms' heterogeneity. Firstly, enterprise productivity within the same industry often substantially differs. Secondly, more productive firms tend to be more involved into international trade and investments than less productive ones. Thirdly, trade liberalization leads to redistribution of market shares between the firms in the industry, more productive firms gaining larger shares. The fourth aspect is that these changes in production shares in the industry lead to growth of the firms' average productivity in this industry even if each firm's productivity remains constant as it is the case in the model by Melitz (2003). Helpman (2005), Drapkin (2010), Kireev (2011) provide detailed review of international trade models with heterogeneous firms.

The work by Melitz (2003) is at the origin of research that takes into account firm heterogeneity. Melitz (2003) developed a dynamic model of general economic equilibrium with one industry. He considers two states of economy, autarky and involvement into international trade, and provides a new insight into the impact of trade on industrial structure and productivity.

The results of the models with heterogeneous firms developed by Bernard et al. (2003) and Yeaple (2002) are in line with the results of Melitz (2003). Firstly, only the most

productive firms working on the domestic market become involved in exporting. Secondly, domestic exporting firms are larger than the firms not involved in exporting. Thirdly, as a result of transition from autarky to open economy, industry level productivity increases as more productive firms extend their activities while less productive ones cannot compete with foreign importers and leave the market.

Following Melitz (2003), Melitz and Ottaviano (2008) study the impact of international trade on countries of different sizes. Okubo (2006), based on the model of vertical linkages in the agglomerations developed by Krugman and Venables (1995) and on the Melitz (2003) approach, shows that when firms' heterogeneity is taken into account, prices on tradable goods do not equalize even under zero trade costs. Less developed country incurs losses from trade liberalization, while better developed country gains. The choice of firms between exporting and FDI is studied by Helpman et al. (2004) and is reviewed in Chapter 5.

Drapkin et al. (2012) extend the model by Melitz (2003). An industry producing an intermediate good is added, following Markusen and Venables (1999). Along with exporting, firms can be involved in FDI as in Helpman et al. (2004). Following Lin and Saggi (2003), backward linkages are studied under international trade and FDI; supply-backward effects are considered. Among the results of the model by Drapkin et al. (2012) are that under transition from autarky to trade, and from trade to FDI, the sign and magnitude of backward linkages depend on the number of national firms, revenue and productivity of average national firm in the initial and resulting equilibrium conditions, which is in line with the existing literature on vertical linkages (Rodrigues-Clare, 1996; Markusen and Venables, 1999; Lin and Saggi, 2003). In case of a substantial technological gap between the firms with FDI and national firms, the sign of backward linkages is ambiguous, unlike in Lin and Saggi (2003). Lin and Saggi (2003) analyze the model with homogeneous firms and conclude that backward spillovers are negative if technological gap is large. Drapkin et al. (2012) conclude that due to vertical linkages, under certain conditions, FDI may stimulate growth of product variety in the industry – FDI recipient and in the supplying industry. The crowding out effect resulting from strengthened competition is greater for the foreign importers than for the national firms. Decrease in intermediate goods prices resulting from FDI stimulates further FDI inflow into the country.

Review of empirical results. There are several types of empirical research on FDI. Firstly, there are case studies for particular countries (for example, for Ireland: Barry (2006)). Secondly, there are industry level studies, usually showing positive dependence between FDI presence and productivity. However, causality between FDI presence and high industrial

productivity is disputable due to potential endogeneity. Besides, MNCs may force less productive domestic firms out of the market, which would also raise average industrial productivity. Overall, in the end of 1990ies - beginning of 2000s, horizontal effects were found to be either insignificant or negative for developing and transition countries, while for developed countries there is some evidence of positive spillovers (Smarzynska-Javorcik, 2004).

Developing and transition countries, having a larger productivity gap compared to the advanced countries, have potentially greater opportunity to gain from FDI; however, evidence shows that firms in the developed countries have larger FDI spillovers, probably due to their higher absorptive capacity (Bode et al., 2009). Bode et al. (2009) also cite the evidence of mutual knowledge spillovers between foreign firms and local firms in the developed countries: an example of spillovers from Japanese investors to local US firms and vise versa. Borenzstein et al. (1998) conclude that FDI contribute to economic growth in developing countries only when a host country has a sufficient absorptive capability defined by human capital. For the US states human capital endowment is also essential in gaining from FDI, according to the evidence cited by Bode et al. (2009). According to Benhabib and Spiegel (1994), human capital affects economic growth, by defining the capacity for innovations of the country and by affecting speed of implementation and spreading of technologies. The same is probably true for the regions and cities within the country.

Hall and Jones (1999) define institutional and political characteristics as 'social capital', saying that it is an important factor in explaining differences in output per worker, while physical capital and educational attainment explain it only partially. 'Social capital', or institutional and political characteristics also affect attraction of FDI and its role for economic growth in the long run. Lipsey (2002) summarizes the studies on FDI effects, from which it follows that local firms benefit from FDI carried out by relatively more productive foreign companies in case technological gap between the local and foreign companies exists but is not too large. Competitive business environment and investment into learning by the local firms is important for FDI spillovers too (Lipsey, 2002). Kathuria (2000) studies Indian manufacturing firms using panel data and concludes that technology transfer effect depends on investment into learning and R&D by local firms. Some of the region specific features, such as 'social capital' or certain tendencies in the firms' behaviour can be controlled for, using regional fixed effects.

Stancik (2007) points out that among the first studies of FDI spillovers using firm level data is that by Aitken and Harrison (1999). They employ a sample containing around 5000 firms in Venezuela for the time period 1976-1989. Their results show that FDI

spillovers on domestic companies with less than 50 workers are positive, while for all domestic companies they are negative, though small. They find that positive spillovers are gained by joint ventures with FDI. They explain the small extent of positive spillovers by not high enough level of development in Venezuela for the time-period under consideration and expect positive effects to become stronger after less productive firms leave the market.

Crozet et al. (2004) provide the evidence that attracting FDI into a certain industry may enhance productivity in this industry. As a result, a country or a region may gain a new comparative advantage. Bode et al. (2009) estimate the aggregate productivity effects of Marshallian externalities from FDI based on the USA states data for 1977-2003. Their contribution to the earlier research on this subject is attention to controlling for Marshallian externalities and other spillovers generated by domestic firms within and outside region in order to differentiate these externalities from FDI spillovers; spatial effects are taken into consideration. Bode et al. (2009) employ a system generalized method of moments (GMM) estimator to deal with potential endogeneity of FDI and the presence of spatial lags. Firms with FDI are found to be a source of positive externalities, while domestic firms - of negative ones.

Among the factors allowing firms to gain from FDI in an industry or in the region, there is local firms' productivity, absorptive capacity, and technological gap compared to foreign firms (Bekes, Kleinert, Toubal, 2009, Damijan et al., 2013). Damijan et al. (2013) studied firms in the transition countries and came to the conclusion that horizontal spillovers are significant, and their importance has been growing during the recent ten years. They also concluded that direct effects and spillovers are affected by geographical distance between a foreign subsidiary and national firms, time period of technology transfer via FDI, absorptive capacity of the national firms, their size, productivity and technological level. According to Damijan et al. (2013), positive horizontal spillovers are equally distributed across size classes in medium or high productivity firms having higher absorptive capacities. Negative horizontal spillovers are typical for smaller firms, and for low to medium productivity firms.

Evidence for transition economies. In the beginning of 2000s foreign firms have dominated in various sectors in CEE countries: they accounted for 73% of capital and 89% of profits in Hungary, in Poland 43% and 66% respectively, in Czech Republic 28% and 92% respectively (Andreff, 2007). According to Andreff (2007), the entry of foreign firms enhances competition in transition countries. The majority of firms, investing in the countries of Central and Eastern Europe, work in capital-intensive sectors, in the sectors, where human capital plays an important role, and in the traditional sectors (labor-intensive ones and sectors intensively using natural resources). In transition economies services sector leads in attraction

of FDI; this way foreign capital facilitates development of this sector. Besides, productivity of firms with FDI is always higher (Marinova and Marinov, 2003). There is empirical evidence that FDI promote technology transfer, particularly, management technologies. As for transfer of tacit knowledge, it does not always take place (Dyker, 2004). FDI promote professional growth of employees, transition from hierarchical management to work in groups, and automation of office work (Andreff, 2007).

Altomonte and Resmini (2001) following Aitken and Harrison (1999) emphasize that the size and sign of FDI effects may vary depending on the characteristics of the country and region, industry, and MNC strategies. Based on Polish data they confirm positive FDI spillovers both through forward and backward linkages. They also conclude that regional specific features may play a role in attracting plants of certain industries to the territory and that government policy should be aimed at sustainable attraction of FDI or at (re)establishing connections between national enterprises so that FDI would serve as a catalyst for economic development.

In the beginning of 2000s the majority of studies were devoted to FDI effects through horizontal spillovers. Smarzynska-Javorcik (2004) studies vertical spillovers. She aims at finding the relationship between presence of FDI in the downstream sector and productivity of domestic firms in the upstream sector, i.e. suppliers of FDI firms. She uses semiparametric estimation method to account for endogeneity of input demand. Based on firm level unbalanced panel data for Lithuania, 1996-2000, she finds positive FDI spillovers in upstream sectors, but no positive horizontal spillovers. She finds that if presence of firms with FDI in the downstream sector increases one-standard-deviation (3% increase in the backward variable), then output of each enterprise in upstream sector raises 10%.

The results show that there are positive spillovers from FDI in the region where they are located and in other parts of the country, although the effects for the other parts of the country are weaker. Foreign firms oriented at local market generate greater effects than those oriented at export. There is no difference between the impact of firms owned only by foreigners and firms with joint domestic and foreign ownership. However, Smarzynska-Javorcik (2004) emphasizes that the positive vertical spillovers may be either due to knowledge spillovers, or due to increased competition among suppliers crowding out less productive firms. The latter can occur if production by domestic firms in the downstream sector decreases, therefore demand for intermediate goods becomes lower, as foreign producers are more efficient and use less intermediate goods, or they purchase intermediate goods abroad. Therefore more research is needed to understand FDI effects and make policy related conclusions.

Based on research using firm level data for 13 countries of Central and Eastern Europe, Rutkowski (2006) concludes that positive FDI spillovers are stronger than crowding out effect and that FDI decrease market concentration by strengthening competition on the host countries' internal markets.

Stancik (2007) analyzes FDI effects on the sales growth rate of domestic companies in the Czech Republic. He studies horizontal and vertical spillovers, both forward and backward, based on enterprise level panel data for 1995–2003. He takes into account the lagged nature of spillovers. The potential endogeneity of FDI related to future industry growth is taken into account. Stancik (2007) finds that domestic companies, especially in upstream sectors, loose from the presence of foreign companies. He finds negative horizontal (2 year lagged) and backward spillovers present mainly in recent years.

Stancik (2007) also provides literature review on empirical research based on data for several countries, including Czech Republic, showing mixed results on FDI spillovers. In the literature on Czech Republic, horizontal effects were found to be of all kinds: positive, negative or insignificant; vertical spillovers proved to be positive in the two studies that he considers. Stancik (2007) assumes that the contradiction can arise partly due to the time period considered. Transition years of 1991-1996 are characterized by mass privatization and unclear ownership structure, and only after 1998 FDI inflow increased significantly.

The papers by Sabirianova et al. (2005) and Hanousek et al. (2012) shed light on direct effects from FDI and the evolution of firms' efficiency. Sabirianova et al. (2005) examine panel data on Czech Republic and Russia for 1992-2000. They find that privatization to domestic owners did not lead to significant improvements in enterprise efficiency. Another finding is that foreign-owned firms are forcing domestic firms out of the top deciles of the efficiency distribution. This process is attributed to slower 'learning' by domestic firms, higher efficiency of new foreign firms, and acquisitions of more efficient domestic firms by foreign investors. Sabirianova et al. (2005) conclude that foreign firms converge to a higher steady state value of efficiency than the domestic firms, and production gap between the foreign and domestic firms in Russia increased during the years 1992-2000; the increase was more pronounced in the first half of transition. Therefore they question the statement that foreign firms' presence helps domestic firms in the host economy to catch up in their efficiency.

A more recent study based on Czech Republic data was done by Hanousek et al. (2012). They employ a stochastic production frontier model and use unbalanced panel data containing over 190 000 firms for the time period 1996-2007. They bring up the dependence between ownership structure and enterprise performance that became evident during the

transition years in the Central and Eastern Europe countries (i.e. the direct FDI effects are considered in this paper). In line with the existing literature, their results show that highly concentrated ownership has positive relation with firm's efficiency. They also find that concentrated foreign ownership is more beneficial than concentrated domestic ownership, demonstrating positive direct FDI effects. They emphasize, though, that a simple majority is not necessarily the best structure for enterprise efficiency improvement. Hanousek et al. (2012) find growing efficiency among the firms in the sample, which leads to the conclusion about improvements in management and corporate governance.

Merlevede and Schoors (2008) study horizontal and vertical FDI spillovers based on panel data for Romanian firms. They separate labor market spillovers from other horizontal spillovers. They also examine the supply-backward linkages arising from firms with FDI over domestic suppliers to domestic firms using domestic inputs. They find that vertical effects are stronger than horizontal spillovers, and that there are supply-backward spillovers in the economy. The spillovers lead to the total factor productivity raise of 20 – 50% during 1998-2001; the effects depend on the domestic company's initial technological level.

Elmas and Degirmen (2009) based on data for Turkey found that technology transfer via FDI, export and import positively affects total factor productivity, FDI providing the major impact. They point out that labor productivity is higher within the foreign firms than within the national ones; firms with FDI are less vulnerable during economic crises; profit to assets ratio and profit to sales ratio are pretty much the same within the firms with FDI and national firms. Based on the same data, Elmas and Degirmen (2009a) analyzed concentration index and did not find connection between FDI and increase in competition in the production sectors. Using panel data and case studies for a number of countries, Henisz (2009) found that if FDI flows are relatively less concentrated, they enhance economic growth and welfare, while relatively more concentrated FDI flows restrain them, as in case of excessive FDI concentration economic policy may be distorted in favor of foreign investors.

Concerning Russia, both positive and negative FDI spillovers were found (Yudaeva et al. (2003), Kadochnikov and Drapkin (2007), Drapkin et al. (2011a). Dolgopyatova (2008) studies the effects of foreign ownership on corporate governance in Russia based on official statistical data, interviews and surveys (in 2005 and 2006) and concludes that FDI induces development of business practices in Russia on the enterprise level and tends to affect government policy.

Drapkin et al. (2011a) using firm level data for the years 1999-2008, find positive horizontal and vertical FDI effects for the Russian domestic firms. The results showed that more productive Russian firms are less affected by foreign companies' competition compared

to less productive firms. However, less productive domestic firms benefit more from vertical effects. In the Russian regions with relatively more diversified economic structure, positive horizontal FDI effects were found, while in regions with narrow specialization, this effect was negative. As for vertical effects, they were found to be positive also in the regions with narrow specialization. In this research we also find that in agglomeration centers, i.e. in cities with highly diversified industrial structure, localization effects from foreign firms' presence (associated with horizontal FDI effects) are positive for national firms. In the other types of cities they are insignificant, and for all national firms producing tradable goods they are negative.

Kadochnikov et al. (2008) examine the impact of regional FDI concentration on productivity of Russian enterprises using firm-level data on medium-sized enterprises for the period 2001-2005 compiled by "Expert-Data" based on Rosstat data; regional and city-level data source is Rosstat. They conclude that FDI concentration leads to positive urbanization effects and to negative localization and competition effects; home market potential is found (market potential is measured as gross regional product per capita and regional population size). It is assumed that positive urbanization effects might arise from input sharing (infrastructure, business services). Our contribution includes diversity index, city-level analysis, and larger database.

Russia's accession to the World Trade Organization (WTO) is a major issue affecting the role of the country in the international trade, and domestic economic situation that is closely interrelated with it. Rutherford and Tarr (2008) develop a comparative static computable general equilibrium model of Russia; they use household data from the Household Budget Survey of Rosstat. They assess the poverty and distributional impact of accession to the WTO on households in seven regions, which are Federal districts (okrugs) of Russia. The model accounts for FDI in business services and endogenous productivity effects generated by additional varieties of business services and goods. Rutherford and Tarr (2008) conclude that these aspects are crucial. National welfare gains are about 4.5% of gross domestic product in the model. However, in a model with constant returns to scale gains are 0.1%. All population groups in all districts are expected to gain from WTO accession. However, in the Northwest region gains are the largest, followed by Far East and Volga regions. In Siberia and the Urals gains are the lowest. Distribution is not changed much within the districts, except for the large gains of the richest decile of the population in the three regions with a lot of FDI. Details of the model are in the previous study by Rutherford et al. (2005).

1.3. Business environment and regional policy in Russia

Institutions and transport infrastructure are considered among the essential factors behind enterprise performance and location choice. Below are some facts on institutions and transport infrastructure in Russia. *Institutional factors.* The role of private investment is important for economic growth and the role of institutions is fundamental for attracting it. Growth of manufacturing sector share and economic diversification would enhance stability in the changing world market situation, taking into account the changes in the oil and gas market trends. Orientation at investment inflows is reflected in the 'accelerated variant' of development in the long term forecast of social and economic development (up to 2030) ratified by the Ministry of Economic Development in 2013[4].

According to the director for macroeconomic research of Higher School of Economics, Sergey Aleksashenko, for economic growth, labour productivity growth is essential, depending in its turn on institutional reforms, such as reforms of judicial system, and investment climate improvement[5]. However, according to the World Justice Project (WJP) Rule of Law Index, currently Russia is on the 92nd place out of 97 countries in Limited government powers indicator; 71st in Absence of corruption; 92nd in Order and security; 83rd in Fundamental rights; 74th in Open government; 68th in Regulatory enforcement; 65th in Civil justice and 78th in Criminal justice[6]. The level of Criminal justice, Limited government powers and Open government (including accessibility and stability of legislation) indicators are lower than in any of BRICS countries.[7] According to the 'Global Entrepreneurship Monitor'[8], in Russia a share of people willing to start a business is 3.8% (in BRICS countries it is on average 21%; in the East European countries – 24%); the number is low due to government policy[9].

Russia is on the 112th place according to the Doing Business index in the annual survey by the World Bank, based on 10 criteria, including the simplicity of connection to the electrical supply network, the number of days and documents needed to get permission for

[4] Vedomosti, (15.05.2013,№82 (3344)) Olga Kuvshinova. The heart of the problem: terminal station. / based on Vladimir Mau 'Economic policy in 2012: between modernization and stagnation' (Russian economy in 2012. Tendencies and perspectives. Moscow. Gaidar Institute. 2013). Interview with the Minister of economic development Andrey Belousov.

[5] Kuvshinova, Olga. Cyprus events will drive Russia into recession. Vedomosti.ru. 22.03.2013. http://www.vedomosti.ru/politics/news/10325671/kipr_obvalit_rossiyu?from=newsletter-editor-choice

[6] The World Justice Project. Russia. http://worldjusticeproject.org/country/russia

[7] Vedomosti, (15.05.2013,№82 (3344)) Olga Kuvshinova. The heart of the problem: terminal station.

[8] Global Entrepreneurship Monitor. 2012 .http://www.gemconsortium.org/docs/download/2645

[9] Vedomosti.ru**,** (17.05.2013) Editorial. What business needs apart from amnesty. http://www.vedomosti.ru/opinion/news/12120521/ne_tolko_amnistiya?from=newsletter-editor-choice

construction and for starting business etc. (the 1st place is attributed to Singapore)[10]. Foreign investors use another indicator of business climate: employment in small and medium sized business and its share in GDP; in Russia this share is around 20%, while in European countries and in USA it is no less than 50%, for comparison[11].

The experts of Higher School of Economics emphasize that freedom, competition and rule of law (that would provide framework for freedom and competition) should be at the center of the new economic model; institutional reforms are essential for economic growth. According to the experts of Higher School of Economics (Evgeny Yasin, Natalia Akindinova, Lev Yakobson and Andrey Yakovlev, 'modifications of legislation, that improve its contents but do not affect rule of law implementation, independence of courts, relations between business and law enforcement institutions, are not sufficient for implementation of economic growth mechanism' [12]

Transport infrastructure. Russia covers around 17 mn square km with 9 time zones[13]. Taking into account the vast territory of Russia, transport infrastructure is especially important. In the State programme – 2020, the Ministry of Transportation writes that transport infrastructure is currently being developed without sufficient coordination. Integrated logistics systems are not implemented; there are no distribution centers, technologies of intermodal transportation are underdeveloped.

According to the director of Institute of Transport Economics at Higher School of Economics Mikhail Blinkin, 80% of the Russian population travel only 2500 km per year. Mobility of the other 20% is much higher, therefore the mean indicator is 6900 km per year; in the EU it is 13000 km per year, according to the data of the Ministry of Transportation and to Eurostat. The losses due to transportation accidents account for 8% of GDP per year. The share of transport costs in the national producers' prices is 15-20% compared to 7-8% in the developed countries[14]. In a region where railroad transportation prices are higher due to insufficient railroad infrastructure, investment is expected to be 1.1-7.5% lower than in a region where transport infrastructure is sufficiently well developed[15]. According to Alexey

[10]Doing Business. World Bank. http://www.doingbusiness.org/methodology/starting-a-business

[11] According to the managing partner of Ernst & Young Aleksandr Ivlev. Vedomosti.ru, (25.06.2013) Mikhail Overchenko, Olga Kuvshinova. Doing Business needs correction. http://www.vedomosti.ru/finance/news/13467661/doing_business_nuzhdaetsya_v_pravke?from=newsletter-editor-choice

[12] Vedomosti.ru, (25.06.2013) Mikhail Overchenko, Olga Kuvshinova. Doing Business needs correction. http://www.vedomosti.ru/finance/news/13467661/doing_business_nuzhdaetsya_v_pravke?from=newsletter-editor-choice

[13] Time zone (Chashvye poyasa Rossii). http://www.timezone.ru/

[14] Vedomosti.ru, (14.05.2013), Margarita Lyutova. How can the economy escape from the infrastructure trap?

[15] Ibid.

Privalovsky (Institute for Geography, Russian Academy of Science), there are no automobile roads and railroads on the 26.5% of Russian land, and on the 23.5% of the land roads density is extremely low[16].

Geography and regional policy in Russia. Among the factors influencing economic development of territories are climate and market capacity. *Climate* in various regions is not explicitly taken into account in this research. However, to provide a broader picture of the country, it should be mentioned that *climate* in Russia is such that energy consumption rate is 3.5 times higher than in Japan, Germany, France, and Great Britain. It is two times higher than in Canada and the USA[17]. As for *market capacity*, in 42 regions of Russia (out of 83) population density is 0.1 – 20 people per square kilometer, which is not sufficient for successful economic development. In 28 regions density is 0.1 – 10 people per square kilometer, which is too low for successful economic development. It seems to be the major reason why the expenses for transport infrastructure development are not repaid quickly. Global competitiveness index for Russia is low and decreasing: the 53rd place in 2005 and the 63rd one in 2011, and certain territories are in especially precarious conditions[18].

Concerning economic development of the regions, in the 1990ies production decreased and regional inequalities strengthened (Granberg, 2001; Shiltsin, 2010). Interregional differences among the majority of regions can be defined as 'quasi-divergence'; the differences among the majority of regions have been decreasing in 1998-2007 by 2.2% per year, but a number of regions-leaders creates an impression of divergence (Shiltsin, 2010). Regional growth factors increasingly differ, although there is some evidence that disparities among the *regions* do not tend to become larger (Russian Economic Report, 2007).

In the framework of Krugman's (1991) model, a territory can become either a periphery, or a center attractive to 'a modern sector' of business. There is international experience of promoting economic growth in the lagging regions, or 'periphery', combining compensating (social) and stimulating (economic) policy. For example, there are Inter-territory Compensation Funds[19] in Spain. Compensating policy in Russia turned out to be insufficient to enhance regional development (Shiltsin, 2010). Among cities, significant differences in economic development are typical both for Russia and for the other countries

[16] Pavel Degtyarev. My native country is too large. Expert-Ural. # 43, 29October – 4 November, 2012, pp. 32-35.
[17] http://serc.carleton.edu/sp/library/google_earth/examples/energy.html
[18] Pavel Degtyarev. My native country is too large. Expert-Ural. # 43, 29October – 4 November, 2012, pp. 32-35.
[19] The official page of Inter-territory Compensation Funds
http://www.dgfc.sgpg.meh.es/sitios/dgfc/en-GB/ipr/oipr/fci/Paginas/inicio.aspx

(Glaeser, Gottlieb, 2009). In this research the accent is made on city level analysis of business performance and location choice.

Policy directed at monotowns. Monotowns represent a special case of urban development. For analytical purposes monotowns[20] (mono-profile or mono-industry towns or districts) are defined as towns specializing in single industry. Their economy is based on one or more town-founding enterprises. There are monotowns – satellites of large cities; monotowns with several town-founding enterprises (of one industry); monotowns with a single town-founding enterprise: a monotown with an enterprise belonging to a holding (either to the main assets or to the secondary assets of a holding) or with an independent enterprise (Monotowns in Russia, 2008; Vorobyev and Scherbinina, 2009).

Specific features of mono-towns depend on the type of economic development of a territory: agglomeration, resource-based or industrial type. Agglomeration zone - center of the European part of Russia - includes enterprises belonging to the chemical industry and metallurgy. Most often it is satellite towns of agglomeration centers. Machine-building enterprises are located in the Central districts partly due to high density of population in this part of a country. Development of monotowns specializing in light industry in the Central region started in the beginning of 19th century; historically and currently it is linked with available work-force and large consumer markets (Vorobyev and Scherbinina, 2009). Resource zone comprises Northern and Eastern territories of the Russian Federation, and the Ural region. Mono-towns specialize on the extraction and initial processing of natural resources: Enterprises belonging to wood and wood processing industries are located in taiga zone (Vorobyev and Scherbinina, 2009). Industrial zone: monotowns are traditionally located in the Ural region and South of Siberia. In Sverdlovsk, Chelyabinsk and Kemerov regions certain territories consist only of monotowns. Economy of monotowns is mainly export-oriented, and thus depends on the situation in the world markets.

[20] The first monotowns appeared as a result of industrial reforms by Peter I (time of reign: 1682 - 1725). Mainly *cloth manufactures and iron-producing* plants. In the beginning of 1930ies monotowns were founded as a part of *territory-producing complexes*); Vorkuta, Norilsk and a number of other towns were built by the GULAG prisoners. During the 2nd World War a lot of monotowns appeared based on evacuated enterprises. During the 1950ies monotowns appeared because new plants were built in small and medium-sized towns, mainly in the Eastern part of Russia (within the program of developing 'less accessible' regions). Within the defense and nuclear policy 'closed' towns with limited circulation of people were created. In the framework of enhancing research and development potential a number of 'research towns' were created. In the beginning of 1990ies the majority of monotowns faced break down of production chains created during the planned economy, decline of production and lack of competitiveness; exception is monotowns specializing on oil and gas. The monotowns, which benefited from the economic growth in 2000-2008 are those specializing in export-oriented industries: oil-extraction, ferrous and non-ferrous metallurgy, chemical industry and the wood industry (Monotowns in Russia, 2008). Due to economies from specialization enterprises located in monotowns can be highly efficient and therefore profitable (Vorobev et al., 2009).

Estimations of the number of monotowns vary among experts. According to the report *Monotowns in Russia* (2008), by the end of 1990is out of 1095 Russian towns around 440 towns and 332 mono-functional 'settlements of the town type' could be classified as mono-towns. They account for 40% of total gross regional product and 25% of population. Their population range from 5-6 thousand people up to 500 thousand people (Tol'yatti). Town-founding enterprises in 145 mono-towns belong to large business groups and companies[21], regional business structures and natural monopolies (Vorobev and Scherbinina, 2009).

A period of stability in monotowns began in 1999 due to growth of prices for monotowns export products on the world market. The monotowns, which profited from the economic growth in 2000-2008 are those specializing in export-oriented industries: oil-extraction, ferrous and non-ferrous metallurgy, chemical industry and the wood industry (Monotowns in Russia, 2008).

Risks associated with situations of crises can be estimated along the following criteria for monotowns: specialization of town economy, share of taxes on physical entities' income, employment level on town-founding enterprises[22], development of small business (Vorobev and Scherbinina, 2009). During the recent economic crisis light industry and machine-building industry experienced significant problems making the towns in the Ural region and Siberia the most vulnerable. Economic situation in monotowns remains unstable due to crisis of 1998 and the recent crisis, and due to reallocation of property between business groups, because monotowns are dependent on the particular industry shocks. Therefore some diversification of their economic structure is necessary. However, due to economies from specialization, enterprises located in monotowns can be highly efficient and therefore profitable. For them diversification mainly concerns departure from specialization on

[21] For example, the share of mining metallurgical industrial complex 'Norilsk nickel' in income of Krasnoyarsk territory (krai) is 50%.

[22] The *concept of monotowns* is closely connected with the phenomenon of town-founding enterprises. Legislation contains several definitions of such type of enterprises. According to the Decree of the government of the Russian Federation from 29 August 1994 № 1001 "On the procedure of defining an enterprise as the '*town-founding*' one and on specific features of selling the indebted enterprises" defines a town-founding enterprise as the one where no less than 30% of the total number of the town (or settlement) workers are employed or that has on its balance objects of social-communal sphere and engineering infrastructure, serving no less than 30% of the inhabitants of the town (or settlement). (Monotowns in Russia, 2008). In the Federal law from 26 October 2002 № 127-FZ "On bankruptcy" those enterprises are considered to be town-founding, which account for no less than 25% of the employed population in the town (Monotowns in Russia, 2008). The town-founding enterprises in the majority of monotowns belong to the wood and wood-processing industries (20%), machinery (17%), food industry (14%), fuel (11%), ferrous and non-ferrous industries (6%); town-founding enterprises in 32% of monotowns belong to the other industries. There are monotowns both in the remote regions and near large cities. Share of workforce employed by a town-founding enterprise can reach 75% (Verkhnyaya Salda, OAO 'Evras Group S.A.') (Monotowns in Russia, 2008).

resource-depending sectors and developing a more advanced production processes for the main product of the monotown leading to a larger added value (Vorobyev and Scherbinina, 2009).

Overall, urban development strategies considered for monotowns in the literature, are further specialization, diversification of the economy, or in the extreme cases, helping people to move to the other cities. To a certain extent, there is a tradeoff between diversity economies and localization economies. The diversification promotion means that a city will develop new industries. However, government can also support an existing industry which means that the policy is localization-oriented. Enterprises from different industries benefit from localization and diversity differently. As a result, the choice of industries for the diversification of monotowns is very important. There is an example of Pittsburg where since the crash of steel industry in 1980ies the city had been successfully developing high-tech sectors of medicine, biotechnology and software.

As for monotown diversification, experts point out *pairs of industries* with the highest probability of joint allocation based on the analysis of industrial diversification in the European countries, such as clothing industry together with textile industry; financial sector and business services sector, etc., using the database of the European Union 'Cluster Observatory'. The types of clusters that can be developed based on monotowns economy are firstly, local clusters – construction, transport, financial sector, food processing industry, business services; secondly, trading clusters – production is exported; thirdly, resource clusters – based on the natural resource endowments (Vorobev, Scherbinina, 2009).

To attract new enterprises and industries in cities and towns, it is needed to understand which of the towns' characteristics are particularly important for enterprises. We analyze the factors influencing enterprise performance in various types of towns, including monotowns, and the determinants of business location choice. The next chapter is devoted to data and methodology used in chapters 3-5.

2. Data description and methodology[23]

2.1. Introduction

The following questions are considered in this chapter. Which aggregation level of OKVED classification allows distinguishing localization and urbanization? Enterprises belonging to which industries are affected by agglomeration in a similar manner? How to measure agglomeration and home market potential in the cities? What types of cities are there in Russia, in terms of different conditions for business? What were the changes in agglomeration and home market potential for different types of cities in 2000s? Which estimation technique is suitable to find the impact of agglomeration and home market potential on enterprise performance?

Several agglomeration indices are considered below: localization index, urbanization index, diversity index (Martin et al., 2011 and Vorobyev et al., 2010), Elison-Glaeser index, Herfindahl–Hirschman index (HHI), Gini index, specialization index (localization quotient). We analyze, which aspects of agglomeration they reflect. A number of strong points and drawbacks of these indices are described. More details on agglomeration indices are in the work by Combes et al. (2008a). Agglomeration is subdivided into localization, urbanization and diversity; details on construction of diversity index are presented. Measures of potential demand and of distances between settlements are considered; their strong points and drawbacks are described. Allocation of economic activity in Russia is briefly described, i.e. where cities and towns are situated and how transport infrastructure is organized. Different types of cities are described: agglomeration centers, cities belonging to agglomeration, monotowns.

The chapter is structured as follows. Section 2.2 is devoted to data description. In Section 2.3 we explain classification of industries into five groups based on their assumed differences in agglomeration effects. Section 2.4 deals with agglomeration indices. In Section 2.5 measurement of home market potential is described. Section 2.6 is devoted to the territory classification. In Section 2.7 estimation of agglomeration economies is described, and Section 2.8 is devoted to the potential issues associated with this estimation. Conclusion follows.

[23] This chapter and chapter 3 are based on the EERC Working paper No10/05 'Spatial concentration and firm performance in Russia' written together with Pavel Vorobyev and Nadezhda Kislyak within the project No R09-0651. The research described in these chapters is performed using a different database; the methods are slightly changed, but the general approach remains the same. In chapters 4 and 5 some of the methods developed in the EERC project are used too.

2.2. Data and descriptive statistics

Research is based on the SPARK-Interfax data augmented with data covering regional and city characteristics. SPARK-Interfax database contains enterprise level data on the organizational form, property, the year of foundation, location, revenue, labour, cost price, profit, and covers the years 1999-2008. The period under consideration covers an advanced stage of transition, after the 1997 new Privatization Law[24] and the 1998 financial crisis. The starting point of transition was the year 1991, the year of dissolution of the Soviet Union and turning from the centrally planned system to the market economy[25].

In the sample constructed based on this data source industries are classified according to 2-, 3-, and 4-digit OKVED classification. The sample used in this project contains 23632 firms, of which 15609 firms belong to the 'tradable' industries, i.e. industries producing tradable goods. The firm level data was filtered so that the majority of indicators are present for the firms in the sample; filtering was based on labor data, as this indicator is the one with the majority of omitted entrances compared to the others. The sample is representative by comparison with complete SPARK databases and with Rosstat data. The database includes 6% firms with foreign ownership. Analysis is based on the balanced panel.

The city level data is taken from Rosstat database on cities with population exceeding 100 thousand people. It covers the years 2002-2008. For the year 2007 there are 12200 firms from SPARK-Interfax database located in these cities, including 8569 firms of tradable industries. Data on city average wage is used in this research. Regional data is collected by Rosstat: gross regional product, the number of patents, transport infrastructure etc. Regional business climate indicators were constructed by the Analytical agency 'Expert' (details on data sources are in Appendix C). The lists of monotowns are taken from the book by Zubarevich (2010) and the report *Monotowns in Russia* (2008).

As there is a tendency for businesses that work in various regions of Russia to register in Moscow, the indicators of localization and diversity in Moscow are probably exaggerated. Researchers sometimes exclude Moscow and St. Petersburg from estimation. However, we did not exclude companies based in Moscow from our models in Chapters 3 and 4, as the results with and without them do not differ significantly; diversity coefficient tends to be more important for enterprises registered in Moscow; localization coefficient – for enterprises

[24] Database of legislation documents 'Consultant Plus'
http://base.consultant.ru/cons/cgi/online.cgi?req=doc;base=LAW;n=13731 (1991)
http://base.consultant.ru/cons/cgi/online.cgi?req=doc;base=LAW;n=28103 (1997)

[25] The first wave of transition in 1991-1995 is associated with price liberalization, mass privatization, financial system reforms, liberalization of external economic activities (Foundations of theory of transition economy, 1996).

registered elsewhere. Overall, the results are robust if companies based in Moscow are excluded from estimation. In Chapter 5 Moscow and St. Petersburg are excluded.

The sample contains manufacturing tradable industries[26]. In the industries included into the sample, there are enterprises that account for around 51% of revenue generated in these industries according to aggregated Rosstat data. Distribution of enterprises among the 5 groups of industries according to the classification presented in this chapter is in Figure 2.1 below.

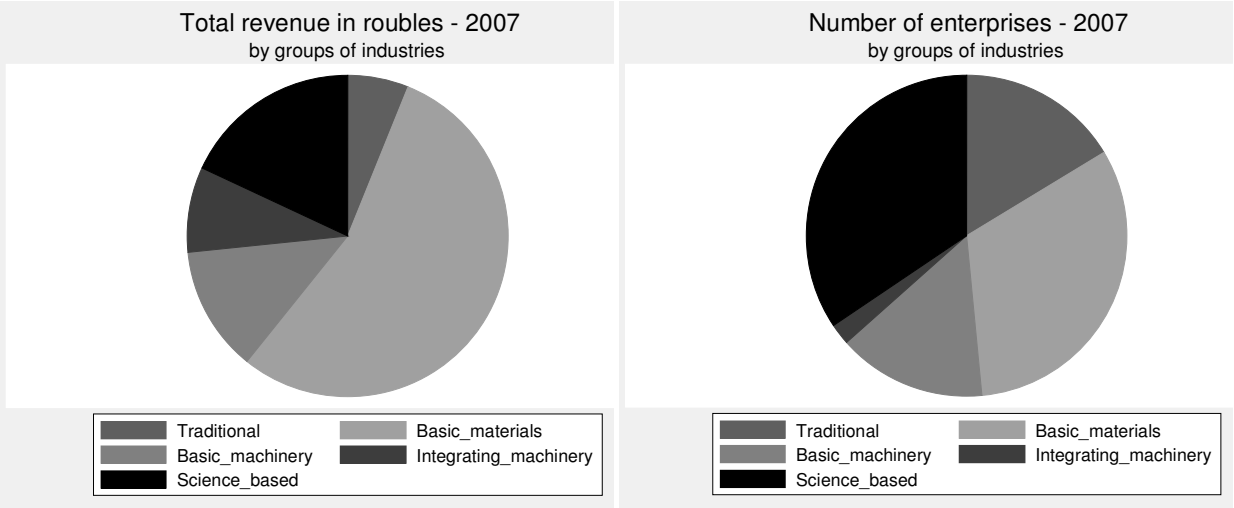


Figure 2.1. Total revenue and shares by industry classes

In terms of revenue, the largest share belongs to the basic materials industries, followed by science based industries, basic machinery, integrating machinery and traditional industries. As for the number of enterprises, enterprises of basic materials industry are the most numerous, followed by science based, traditional, basic machinery industries, and integrating machinery industry. This distribution is similar to that in the enterprise level database for Russia covering the years 2001-2005, which was used by Vorobyev et al. (2010).

Table 2.1. Average profit of firms, by type of city and class of industry, mn roubles, 1999-2008

City type	Traditional	Basic materials	Basic machinery	Integrating machinery	Science based	All tradables
Agglomeration center	13.45	42.60	13.36	24.09	17.21	23.22
Within agglomeration; not center	5.19	29.70	16.39	87.89	10.52	19.60
Monotown within agglomeration	0.67	309.61	0.73	416.98	1.27	190.35

[26] SPARK: OKVED 10 (CA) – 37(DN), 40(E), 41(E) – i.e.whole E, 45(whole F), 60(part of I), 65(part of J), K and 'high tech' industries

City type	Traditional	Basic materials	Basic machinery	Integrating machinery	Science based	All tradables
Monotown outside agglomeration	1.45	169.04	36.40	38.58	-5.73	93.76
Other cities	5.17	34.36	6.85	10.71	4.97	15.21
Total	7.12	48.57	11.57	25.51	11.63	22.95

Based on the data used in our research, when all industries, both tradable and not, are taken into account, average profit in agglomeration centers over the years 1999-2008 is 36.59 mn roubles, while in all cities and towns on average it is 29.75 mn roubles. The same time, if only tradable industries are taken into account, beginning with the year 2004 profit is persistently higher in the cities other than agglomeration centers, and on average over the years 1999-2008 it is 23.22 mn roubles in the agglomeration centers (Table 2.1 above) and 22.95 mn roubles in all cities and towns on average. Data used by Vorobyev et al. (2010) also shows that average profit is higher in the 17 large cities, than in the other towns – 22 versus 19 mn roubles on average for the time-period 2001-2005.

Table 2.2. Average revenue of firms, by type of city and class of industry, mn roubles, 1999-2008

City type	Traditional	Basic materials	Basic machinery	Integrating machinery	Science based	All tradables
Agglomeration center	227.12	581.41	318.95	977.16	242.33	353.60
Within agglomeration; not center	162.08	432.23	410.38	3768.99	210.22	399.72
Monotown within agglomeration	20.82	3187.47	91.67	5093.59	33.85	2036.75
Monotown outside agglomeration	43.14	1848.56	543.06	1527.64	1283.63	1155.91
Other cities	123.43	413.21	246.96	935.04	103.36	253.85
Total	147.77	592.73	299.17	1207.97	187.58	349.69

As it would be expected, profit and revenue are higher in agglomeration centers than in the other towns and cities, both within and outside agglomerations, except for monotowns. Tables 2.1 and 2.2 show that in monotowns enterprises have on average the highest profit and the highest revenue. However, this pattern varies among the industries. Concerning revenue, it is in basic materials and integrating machinery industries that enterprises located in monotowns have revenues higher than enterprises in the other types of towns.

Table 2.3. Average number of employees, by type of city and class of industry, 1999-2008

City type	Traditional	Basic materials	Basic machinery	Integrating machinery	Science based	All tradables
Agglomeration center	246.87	487.29	651.08	815.53	344.83	418.00
Within agglomeration; not center	318.57	572.33	507.54	1140.03	375.55	478.79
Monotown within agglomeration	114.30	1220.39	313.57	7489.68	108.75	1120.94
Monotown outside agglomeration	133.69	1322.91	862.76	1429.14	2152.89	1020.70
Other cities	255.40	514.47	616.43	1575.68	311.63	448.49
Total	249.59	574.87	637.90	1365.80	341.72	468.92

Table 2.4. Average fixed assets, by type of city and class of industry, 1999-2008

City type	Traditional	Basic materials	Basic machinery	Integrating machinery	Science based	All tradables
Agglomeration center	59.95	121.62	73.84	174.60	52.45	76.74
Within agglomeration; not center	41.54	142.22	46.64	248.26	60.57	91.46
Monotown within agglomeration	4.24	506.97	23.36	703.43	4.06	319.03
Monotown outside agglomeration	12.25	489.48	114.52	313.86	333.18	297.30
Other cities	36.37	127.33	155.75	229.37	59.52	98.10
Total	41.34	156.08	120.98	224.12	57.29	100.20

Tables 2.3 and 2.4 show that on average for all tradables, enterprises located in monotowns employ the largest amount of inputs (workers and fixed assets). Across industries, the largest firms on average in terms of inputs belong to the integrating machinery sector.

Table 2.5. Productivity[27] by type of city and class of industry, 1999-2008

City type	Traditional	Basic materials	Basic machinery	Integrating machinery	Science based	All tradables
Agglomeration center	0.64	1.29	0.82	1.29	0.79	0.92
Within agglomeration; not center	0.39	0.77	0.60	1.41	0.40	0.59
Monotown within agglomeration	0.16	1.23	0.31	0.75	0.23	0.83
Monotown outside agglomeration	0.23	0.73	0.57	1.06	0.35	0.59
Other cities	0.34	0.73	0.46	0.62	0.30	0.49
Total	0.41	0.91	0.58	0.90	0.56	0.66

[27] Revenue per one employee, mn roubles.

As it would be expected, Table 2.5 shows that enterprises located in agglomeration centers have on average the highest productivity. Monotowns located within agglomerations have high productivity on average as well. However, productivity varies substantially among industries, and difference in productivity among cities depends on the industries typical for the cities. Concerning differences in productivity in a certain industry among cities, for example, in basic materials and integrating machinery industries monotowns have enterprises with relatively high productivity. Table 2.6 below contains descriptive statistics for the main variables used in econometric analysis in the following chapters.

Table 2.6. Descriptive statistics of variables, 1999-2008

Variable		(1) **Traditional**	(2) **Basic materials**	(3) **Basic machinery**	(4) **Integrating machinery**	(5) **Science based**	(6) **All tradables**
Logarithm of enterprise revenue	**Mean**[28]	**16.47**	**17.42**	**17.90**	**18.61**	**16.77**	**17.14**
	Std.dev	2.25	2.57	1.93	2.17	2.29	2.39
Logarithm of fixed assets	**Mean**	**14.92**	**15.80**	**16.13**	**17.10**	**15.12**	**15.51**
	Std.dev	2.49	2.78	2.23	2.39	2.73	2.68
Logarithm of labour	**Mean**	**4.55**	**5.00**	**5.52**	**6.01**	**4.59**	**4.89**
	Std.dev	1.45	1.66	1.37	1.59	1.59	1.60
Logarithm of localization coefficient (lloc) 3-digit	**Mean**	**8.42**	**10.99**	**11.47**	**11.80**	**18.31**	**13.25**
	Std.dev	9.34	9.76	9.75	9.75	7.78	9.83
A share of industry (3-digit) in total city revenue (core)	**Mean**	**.05**	**.078**	**.066**	**.16**	**.06**	**.07**
	Std.dev	.15	.19	.15	.25	.12	.16
Localization quotient	**Mean**	**47.78**	**21.18**	**19.395**	**29.81**	**6.20**	**20.28**
	Std.dev	333.69	101.81	79.92	77.14	56.71	154.55
Diversity coefficient (divpq)	**Mean**	**.26**	**.294**	**.302**	**.27**	**.44**	**.34**
	Std.dev	.25	.26	.24	.23	.29	.27
Logarithm of GRP (lgrp)	**Mean**	**12.08**	**12.24**	**12.14**	**12.02**	**12.72**	**12.36**
	Std.dev	1.4	1.45	1.37	1.30	1.61	1.51
Logarithm of urbanization coefficient (lurbpq)	**Mean**	**22.71**	**23.17**	**23.46**	**22.92**	**24.55**	**23.61**
	Std.dev	3.23	3.21	2.72	3.17	2.91	3.12
Dummy	**Mean**	**.0006**	**.014**	**.007**	**.022**	**.001**	**.007**

[28] Average for the years 1999-2008.

Variable		(1) Traditional	(2) Basic materials	(3) Basic machinery	(4) Integrating machinery	(5) Science based	(6) All tradables
variable for a city-forming enterprise	Std.dev	.021	.12	.08	.15	.033	.08
Dummy variable for a monotown[29]	**Mean**	**.057**	**.08**	**.082**	**.07**	**.006**	**.05**
	Std.dev	.23	.27	.28	.26	.078	.22

2.3. Industrial classification

As it was mentioned above, according to the life cycle theory (Vernon, 1960), new industries benefit from location in diversified places. It leads to the idea that agglomeration effects should differ across industries. We have adopted the approach to consider specific groups of industries from Henderson (2003), who analyzes two groups of industries – machinery and high tech. Based on Pavitt taxonomy of manufacturing industries (Pavitt, 1984) we have disaggregated our sample of firms into five industrial groups (Table 2.7)[30].

[29] Monotowns of large business (Monotowns in Russia, 2008)

[30] The list of 3-digit industries forming each group is presented in the working paper by Vorobyev et al. (2010), available on the eerc web site: http://www.eerc.ru/paper?page=8&sort=theme&type=asc

Table 2.7. Industrial classification for the empirical analysis of agglomeration economies

№	Classification used in this research	Analogue in Pavitt (1984)	Description	3-digit industries: examples
1	Traditional goods	Traditional goods	Industries producing goods that have been produced since a long time ago; mainly consumer goods	Manufacture of made-up textile articles; manufacture of leather clothes; manufacture of products of wood, etc.
2	Basic materials	Scale intensive	Mainly industries with high returns to scale, making homogeneous intermediate goods	Textile weaving; manufacture of bricks, tiles and construction products, manufacture of cast iron and steel, etc.
3	Basic machinery	Specialized suppliers	Industries producing machinery not included into group 4	Manufacture of agricultural and forestry machinery; manufacture of machine tools, etc.
4	Integrating machinery	Specialized suppliers	Integrating sub-industries of machinery, which make complicated products, consisting of numerous spare parts	Manufacture of motor vehicles; manufacture of motorcycles and bicycles, etc.
5	Science based	Science based	Industries, which actively use innovations and new knowledge	Processing of nuclear fuel; Manufacture of aircraft and spacecraft; software consultancy and supply, etc.

Source: Vorobyev et al. (2010), based on Pavitt (1984)

We assume that agglomeration economies for *traditional goods* and *basic materials* arise primarily from localization, and microfoundations behind these economies are sharing and matching (see Table 2.8). Sharing means that firms can decrease their costs by using common resources and benefit from higher variety of intermediary suppliers. It is a very important source of economies for traditional goods producers because their products are differentiated, and therefore they need suppliers of specific intermediary products. When goods become more complicated, the role of learning externalities strengthens, implying the increasingly important role of diversity economies. It is relevant for basic machinery and science based industries; for basic materials diversity economies play a much smaller role. Science based industries also benefit from industrial localization because it facilitates exchange of ideas, infrastructure sharing and attracts specialized highly qualified human capital (see Table 2.8).

Table 2.8. Microfoundations behind agglomeration economies for different groups of industries

	Sharing	Matching	Learning
Traditional goods	Significant, arise from localization	Significant, arise from localization	Not important
Basic materials	Significant, arise from localization	Significant, arise from localization	Not important
Basic machinery	Important; arise both from localization and from diversity	Important; arise from localization	Important; arise mainly from diversity
Integrating machinery	Important; arise both from localization and from diversity	Very important; arise from diversity	Important; arise mainly from diversity
Science based	Important; arise both from localization and from diversity	Important; arise both from localization and from diversity	Very important; arise mainly from diversity

Source: Vorobyev et al. (2010)

Traditional goods producers have medium labor costs associated with training; that is, they can easily train workers who have no special skills. At the same time, these producers have no large gains from variety; that is, they do not need a mix of qualifications. These industries can be efficient both in diversified cities and in monotowns. Producers of basic materials are quite similar to traditional goods producers. Localization economies allow realizing potential returns to scale as external returns. But in Russia firms producing basic materials have very large scale, often playing a role of the city-forming enterprise in a monotown, and do not need any additional localization.

For basic machinery availability of specialized labor and diversity of suppliers are very important, and these producers tend to benefit both from localization and diversity economies. However, in Russia machinery industries experience a very deep crisis associated with low competitiveness in situation of so called Dutch Disease. Performance of enterprise, including its ability to benefit from agglomeration economies is likely to depend primarily on management skills and success of modernization programs. Integrating machinery is similar to basic machinery, but it depends greater on the returns to scale and availability of specialized suppliers. In Russia integrating machinery is represented by large city-forming enterprises such as Avtovaz in Tolyatti.

Finally, science-based (high-tech) industries need very qualified labor in various research areas. They benefit from higher diversity and accumulation of human capital in large cities such as Moscow, Saint-Petersburg, Novosibirsk and Yekaterinburg. At the same time, in Russia there are a lot of small cities specializing in science based industries (in the first place, nuclear). They were built during the Soviet times and reflect localization policy of central planning authorities.

Illustration of relative importance of localization and diversity economies for different groups of industries is presented in Figure 2.2 below.

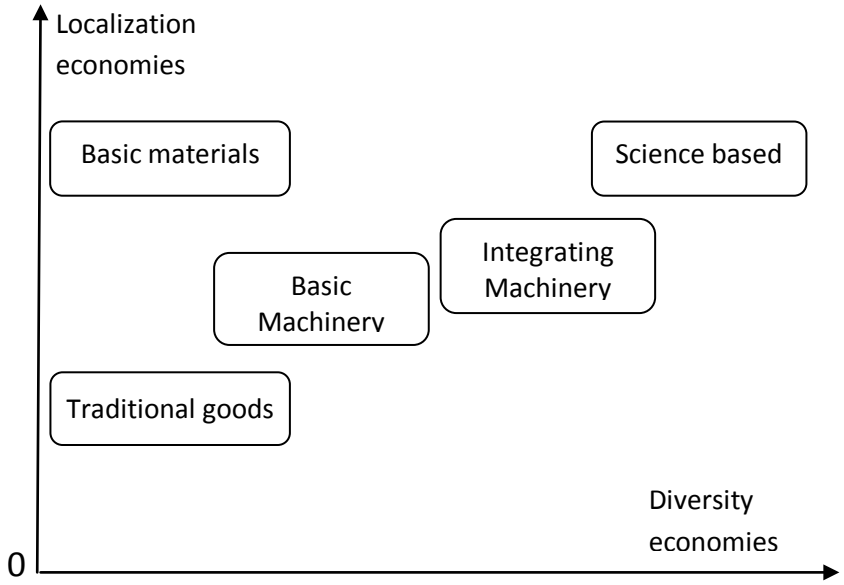


Figure 2.2. Localization and diversity economies for different classes of industries

Source: Vorobyev et al. (2010)

At the enterprise level, diversity effects are usually found to be significant for new[31] enterprises, while localization externalities are important for mature enterprises with standardized manufacturing production (Duranton and Puga, 2001; Neffke et al., 2008). In Tables 3.9 and 3.10 we compare agglomeration economies for the 'old' and 'young' firms in Russia. It reflects another aspect of sensitivity to agglomeration effects: enterprises created during planned economy ('old enterprises') vs. enterprises established after privatization ('young enterprises'), i.e. the enterprises that had a possibility to choose their location. For both types of enterprises all agglomeration effects are found to be significant, however, for 'young' enterprises they are stronger.

2.4. Agglomeration indices

Measuring agglomeration. An important question in empirical studies is a choice of an appropriate measure to capture the extent of agglomeration. Behind each indicator there is a certain mechanism. For example, the number of plants included into analysis allows capturing the possibility of knowledge circulation between firms (Henderson, 2003).

Measurement of localization scale. For the analysis of localization we suggest using the concepts of the *level of economic activity* in a city and *specialization* ('inequality' between

[31] Founded starting with the year 1995 (end of the check privatization).

industries). Following Vorobyev et al. (2010), we suggest the following indicator for measuring localization as the *level of economic activity*:

$$loc_{it}^{jz} = \ln(revenue_{t}^{jz} - revenue_{it}^{jz} + 1), \tag{2.1}$$

where $revenue_{t}^{jz}$ – the revenue of all firms belonging to an industry *j* and located in a city *z*; $revenue_{it}^{jz}$ – revenue of a firm *i* belonging to an industry *j* and located in a city *z; t* is time. This indicator, although with labor instead of revenue, was used by Martin et al. (2011). After comparing indices based on labour and on revenue, the approach with revenue was chosen. Labour productivity is known to be comparatively low in Russia; some enterprises still have a tendency to employ a lot of workers rather than modernize production process. Therefore the indices with revenue seem to reflect agglomeration more objectively.

Localization is measured in the literature based on various industry aggregation levels. Vorobyev et al. (2010) constructed regression with localization based on 3-digit, 2-digit aggregation and on classification into 5 classes of industries developed by them. 3-digit aggregation level is basic in the literature (Henderson, 2003; Vakhitov, 2008; Beaudry and Schifaerova, 2009). Experimenting with different levels of aggregation, the researchers have assessed industrial scope of localization economies. For instance, if significant localization economies were found on the 3-digit level, but not 2-digit level, it would mean that localization economies arise only from interaction between very similar firms. It helps clarify the mechanisms behind localization economies. We estimate coefficients on 3-digit level, as it allows distinguishing localization and diversity relatively more clearly: at 2-digit level and at the level of industry classes it is already less evident if it is still localization effects that matter or it is diversity among industries that starts playing a role. In Tables 2.9 and 2.10 below is descriptive statistics of localization coefficient for types of cities and groups of industries.

Table 2.9. Descriptive statistics for localization, different types of cities

City type	mean(lloc)	sd(lloc)	min(lloc)	max(lloc)	N(lloc)	median(lloc)
Agglomeration center	20.38	5.82	0	26.68	52,955	21.59
Within agglomeration; not center	8.02	9.22	0	23.76	9,012	0
Monotown within agglomeration	8.20	9.74	0	23.48	679	0
Monotown outside agglomeration	9.49	9.92	0	26.62	10,084	2.34
Other cities	9.82	9.41	0	25.55	96,557	14.11

Table 2.10. Mean of ln(localization) across types of cities and classes of industries

City type\class of industry	Traditional	Basic materials	Basic machinery	Integrating machinery	Science based
Agglomeration center	16.71	18.68	18.29	16.43	22.48
Within agglomeration; not center	3.20	5.71	6.33	7.19	13.44
Monotown within agglomeration	0	10.26	9.68	21.38	11.59
Monotown outside agglomeration	2.53	7.23	8.67	13.26	4.69
Other cities	5.92	7.40	8.43	9.07	13.79

To measure *specialization* we use localization quotient (De Siano and D'Uva, 2010; Glaeser et al., 1992; Henderson et al., 1995):

$$LQ_t^{jz} = \frac{revenue_t^{jz} / revenue_t^{z}}{revenue_t^{j} / revenue_t}, \quad (2.2)$$

where $revenue_t^z$ – revenue of all industries in a city z, $revenue_t^j$ – revenue of an industry j in a country, $revenue_t$ – revenue of all industries in a country; t is time.

The concepts of the *level of economic activity* and *specialization* reflect different types of agglomeration. The level of economic activity is a good indicator of agglomeration leading to close interaction between firms. Specialization indices allow comparing development of various industries in different geographical locations such as the cities and therefore helps explain trade flows between them. The first type of indices is more appropriate for this research. As for localization quotient, it is useful for understanding concentration of various industries in various cities. In Table 2.11 below there is descriptive statistics of localization quotient (LQ): average values of LQ for various groups of industries in various types of cities.

Table 2.11. Localization quotient across city types and classes of industries

City type	Traditional	Basic materials	Basic machinery	Integrating machinery	Science based	Total
Agglomeration center	1.97	1.97	1.79	3.58	2.72	2.23
Within agglomeration; not center	74.16	51.95	22.73	35.81	7.71	28.18
Monotown within agglomeration	4.39	23.52	7.53	52.67	1.42	13.34
Monotown outside agglomeration	38.08	26.90	22.36	84.75	105.26	19.00
Other cities	65.35	27.63	28.48	37.34	9.06	25.19
Total	47.78	21.18	19.40	29.81	6.20	17.98

Table 2.11 above shows that concentration of various industries in agglomeration centers is the lowest, compared to its concentration in the country on average. Then follow monotowns, and then other cities. As for classes of industries, concentration is the highest for traditional industries, then follow integrating machinery, basic materials, basic machinery and science based industries.

Diversity can be reflected by such indicators as industry employment, total urban area population, total local employment, the number of industries in a city. Among the indices that measure diversity are Herfindahl–Hirschman index, Gini index of diversity, Theil index (Combes et al., 2008a), Ellison-Glaeser index (Elison and Glaeser, 1997) and the Duranton and Overman index (continuous approach). Combes et al. (2008a) provide an overview of approaches to measuring diversity. Out of the last two indices we have chosen Elison-Glaeser index based on its computational convenience; the index is explained below. Besides, a new approach to measuring diversity is proposed based on the paper by Vorobyev et al. (2010).

Elison-Glaeser index reflects both geographical concentration of an industry and concentration of activities within the industry (Combes et al., 2008a). The original index by Elison and Glaeser was designed to measure industrial concentration. Several modifications of this index were developed in the literature. Vakhitov (2008) uses the following index:

$$\gamma = \frac{G - H}{1 - H}$$

Here H – Herfindahl index of industrial concentration:

$$H_t^{jz} = \sum_{i \in S_i^z} \left(\frac{revenue_{it}^{jz}}{revenue_t^{jz}} \right)^2$$

where $revenue_{it}^{jz}$ - revenue of a firm *i* belonging to an industry *j* and located in a city *z*; $revenue_t^{jz}$ – revenue of firms belonging to an industry *j* in a city *z; t* is time.

G – Gini index showing 'raw' geographic concentration:

$$G = \frac{\sum_{z=1}^{Z} \left(s_z - x_z \right)^2}{1 - \sum_{z=1}^{Z} x_z^2},$$

where the number of cities is *z*; $s_z = revenue_{jz} / revenue_j$ - a share of industry located in a city in the overall output of the industry; $x_z = revenue_z / revenue$ is a share of a city in the output of a country. The index γ allows comparing concentration among various industries, and the actual distribution of economic activity is compared to the uniform distribution.

Two variations of this index, for cities and for industries (Henderson, 2003), were calculated for the Russian cities and industries; the results are presented in Tables 2.12 - 2.14

below. Henderson (2003) offered a modification of Elison-Glaeser index applied to measure *diversity* in city, i.e. lack of local diversity or the degree of local specialization. Following his approach we measure non-diversity of revenue *in a city*:

$$S_z(t) = \sum_j \left(\frac{revenue_{jz}(t)}{revenue_z(t)} - \frac{revenue_j(t)}{revenue(t)} \right)^2,$$

where $revenue_{jz}(t)$ – revenue of firms belonging to an industry *j* in a city *z*; $revenue_z(t)$ – total revenue of firms in a city *z*; $revenue_j(t)$ – revenue of firms belonging to an industry *j* in all cities of a country; $revenue(t)$ – total revenue of firms in a national economy; *t* is time.

$S_z(t)$ is a sum of squared deviations of industry *j*'s share in a city *z* of local relevant revenue from industry *j*'s national share. $S_z(t)$ ranges from zero, when city's share over all industries is the same as national shares so the city is perfectly diverse, to two in the limit in case of high concentration. For example, if one industry is extremely concentrated in the city but not in the country, while another industry is extremely concentrated in the country but not in the city, the value of the index strives to 2.

Table 2.12. Elison-Glaeser index, the version for cities, calculated for types of cities

City type	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	Total
Agglomeration center	0.08	0.07	0.05	0.05	0.05	0.05	0.06	0.07	0.07	0.07	0.06
Within agglomeration; not center	0.29	0.29	0.27	0.28	0.26	0.27	0.28	0.31	0.32	0.33	0.29
Monotown within agglomeration	0.41	0.46	0.45	0.49	0.44	0.46	0.53	0.55	0.55	0.55	0.49
Monotown outside agglomeration	0.41	0.43	0.41	0.38	0.37	0.39	0.39	0.42	0.43	0.44	0.41
Other cities	0.30	0.29	0.28	0.27	0.26	0.27	0.28	0.29	0.30	0.31	0.29
Total	0.24	0.24	0.22	0.21	0.21	0.21	0.22	0.23	0.24	0.25	0.23

In Table 2.12 above, as it was expected, Elison-Glaeser index for cities shows that industrial structure in the agglomeration centers is the most diverse; the highest level of specialization is typical for monotowns, and the other cities within and outside agglomeration have similar levels of concentration. Over the years, diversity of agglomeration centers initially slightly increased, then slightly decreased, being 0.08 in 1999 and 0.07 in 2008. In monotowns and the other towns specialization slightly increased, with the highest increase in the monotowns within agglomerations. Overall, for all city types diversity stayed almost the same, 0.24 in 1999 and 0.25 in 2008.

According to the analysis by Vorobyev et al. (2010) based on 2001-2004 data economy is more diversified (less concentrated) in the larger cities (Elison-Glaeser index is 0.03 both for 2- and 3-digit classifications), and more concentrated in the smaller towns (Elison-Glaeser index is 0.64 for 2-digit classification and 0.62 for 3-digit). In Table 2.13 below, the details on 17 agglomeration centers are provided.

Table 2.13. **Elison-Glaeser index for major cities ('agglomeration centers')**

City	1999	2004	2006	2008
Moscow	.033	.040	.060	.057
St. Petersburg	.057	.021	.023	.034
Samara	.173	.027	.028	.066
Ekaterinburg	.149	.061	.031	.023
Nizhny Novgorod	.110	.091	.117	.103
Rostov-on-Don	.092	.053	.056	.096
Novosibirsk	.208	.047	.105	.116
Omsk	.197	.047	.115	.048
Chelyabinsk	.093	.069	.078	.072
Kazan'	.103	.146	.159	.102
Krasnoyarsk	.259	.166	.229	.259
Irkutsk	.097	.054	.052	.062
Vladivostok	.094	.079	.062	.067
Perm'	.071	.088	.167	.212
Ufa	.157	.114	.110	.102
Kaliningrad	.068	.049	.055	.094
Vologda	.074	.036	.037	.041

Elison-Glaeser index for measuring concentration *within industry:*

$$S_j(t) = \sum_z \left(\frac{revenue_{jz}(t)}{revenue_j(t)} - \frac{revenue_z(t)}{revenue(t)} \right)^2,$$

where $revenue_{jz}(t)$ – revenue of firms belonging to an industry *j* in a city *z*; $revenue_z(t)$ – total revenue of firms in a city *z*; $revenue_j(t)$ – revenue of firms belonging to an industry *j* in all cities of a country; $revenue(t)$ – total revenue of firms in a national economy*; t* is time.

$S_j(t)$ is a sum of squared deviations over cities of each city *z* share in revenue of industry *j* from city *z* share in national revenue. $S_j(t)$ ranges from zero, when industrial share in a city is the same as industrial share over all cities, so the industry is not concentrated in particular cities but equally distributed in the country, to two in the limit in case of total concentration.

Table 2.14. Elison-Glaeser index for industries, for classes of industries

	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	Total
Traditional	0.10	0.09	0.09	0.09	0.09	0.10	0.10	0.09	0.09	0.10	0.09
Basic_materials	0.17	0.18	0.14	0.13	0.12	0.13	0.12	0.14	0.14	0.15	0.14
Basic_machinery	0.25	0.18	0.10	0.08	0.08	0.07	0.07	0.07	0.08	0.09	0.10
Integrating_ machinery	0.26	0.29	0.11	0.09	0.23	0.13	0.16	0.16	0.13	0.22	0.18
Science_based	0.17	0.14	0.15	0.17	0.16	0.17	0.17	0.16	0.15	0.14	0.16
Total	0.13	0.12	0.10	0.09	0.09	0.09	0.09	0.10	0.09	0.10	0.10

Urbanization coefficient measures total revenue of firms belonging to all industries in a city, except for the industry under consideration:

$$\ln\left(\text{urb}\right)_t^z = \ln(revenue_t^z - revenue_t^{jz} + 1) \quad (2.3)$$

where $revenue_t^{jz}$ – revenue of all firms belonging to an industry *j* and located in a city *z;* $revenue_t^z$ - revenue of all firms in a city *z; t* is time. As it was mentioned above, *localization and urbanization* indices are based on revenue; revenue is measured in nominal prices. The distribution of average localization and urbanization levels among the types of cities remained relatively similar over the years 1999-2008. Means for these years are shown in Table 2.15.

Table 2.15. Average localization and urbanization levels, for types of cities, average for 1999-2008[32]

City type	Mean(lloc)	Mean(lurbpq)
Agglomeration center	20.35	26.14
Within agglomeration; not center	8.03	21.64
Monotown within agglomeration	8.23	22.22
Monotown outside agglomeration	9.55	22.54
Other cities	10.19	21.88

One of the ways to measure Jacobs externalities is by *Herfindahl index of diversity*:

$$HHI_t^{jz} = \sum_{k \neq j} \left(\frac{revenue_t^{kz}}{revenue_t^z - revenue_t^{jz}} \right)^2$$

where $revenue_t^{kz}$ – revenue of firms belonging to industry *k* in a city *z*; $revenue_t^z - revenue_t^{jz}$ – revenue of the firms belonging to all industries except for industry *j* in city *z*; *t* is time.

$dvf_t^{jz} = \ln\left(\frac{1}{HHI_t^{jz}}\right)$ reflects the diversity faced by firms of industry *j* in city *z* at time *t* (Martin et al., 2011). This index seems not to capture economic structure of the territory, as it is not clear if there is only one, or many industries other than the industry *j* in an city *z*.

[32] In logarithms, for all industries present in the database

Diversity index. Following the approach of Vorobyev et al. (2010), in this research diversity is analyzed in two dimensions. The first one, *variety,* shows how many different industries exist in a city. The second one, *inequality* indicator, reflects how evenly they are distributed. It seems that the more evenly the economic activity is distributed among industries the more diverse economy of a city is.

A typical problem for diversity indices consists in their instability to elimination of the dominant industry. Let us imagine 2 cities. In the first city 5 industries with revenues (10, 1, 1, 1, 1) are located, and in the second city there are 4 out of 5 industries with revenues (0, 1, 1, 1, 1). The first city demonstrates very high concentration in the first industry, which accounts for 70% of revenue. However, the second city has equal distribution. A typical index of inequality shows that a city (0, 1, 1, 1, 1) is more diversified than a city (10, 1, 1, 1, 1) which is incorrect. Such indices as classical versions of Herfindahl-Hirschman (HHI), and Gini (Combes et al., 2008a) measure inequality. Vorobyev et al. (2010) suggested a new indicator of diversity, which takes into account both variety and inequality:

$$div_t^z = \sum_{j=1}^{s} \left(\frac{revenue_t^{jz}}{revenue_t^z} \right)^{\frac{1}{s}}, \tag{2.4}$$

where $revenue_t^{jz}$ – revenue of the industry *j* in a city *z*; $revenue_t^z$ – revenue of all industries in a city *z*; *s* – total number of industries in a city; *t* is time; in our study diversity index is calculated based on 3-digit industrial classification.

The index was normalized in order to receive the values ranging from 0 to 1. If all industries in a city have equal revenue then the diversity index ('DIV') takes on the value

$$div_t^z = \sum_{j=1}^{s} \left(\frac{1}{s} \right)^{\frac{1}{s}} = s^{1-\frac{1}{s}} \approx s. \tag{2.5}$$

If the entire revenue is concentrated in one industry, the index DIV takes on the value

$$div_t^z = 1^{\frac{1}{s}} + \sum_{j=2}^{s} 0^{\frac{1}{s}} = 1. \tag{2.6}$$

The normalized coefficient is given by the following formula

$$div_t^z = \frac{\sum_{j=1}^{s} \left(\frac{revenue_t^{jz}}{revenue_t^z} \right)^{\frac{1}{s}} - 1}{s^{1-\frac{1}{s}} - 1}. \tag{2.7}$$

This coefficient takes on the value 1 if the distribution of revenue among industries is perfectly equal, and the value 0 if all revenue is concentrated in one industry. Comparison of the indices DIV, HHI and Gini for 3 industries in a city is presented in the Figure 2.3 (left-

hand chart). Vorobyev et al. (2010) suggested the following version of the index with s being the number of industries in the city:

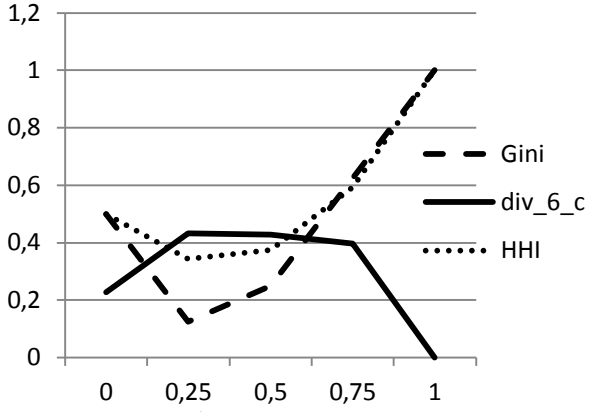
$$div_t^z = \frac{\sum_{j=1}^{s}\left(\frac{revenue_t^{jz}}{revenue_t^z}\right)^{\frac{1}{s}} - 1}{s-1}. \quad (2.8)$$

With equation (2.7) and *s* being the number of industries in a city, as in Vorobyev et al. (2010), the problem is that the index cannot reflect changes in the variety of industries. Under perfectly equal distribution the index is 1 regardless of number of industries in a city. It is due to the normalization procedure. To introduce variety into the index its value with the number of industries striving to infinity was calculated:

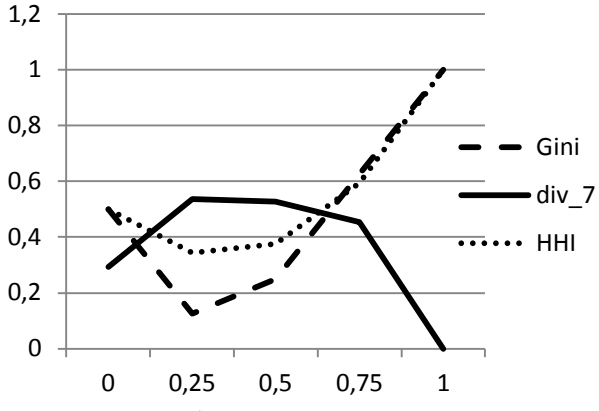
$$div_t^z = \sum_{j=1}^{s}\left(\frac{1}{s}\right)^{\frac{1}{s}} = s^{1-\frac{1}{s}} \underset{s\to\infty}{\to} s. \quad (2.9)$$

Then $s^{1-\frac{1}{s}}$ was substituted by *s* in equation (2.7) and the expression (2.8) was received. In this work, another version of this index is used, with *s* being the number of industries in a country. Variety is reflected by comparing industrial structure in the city and in the country, and it is not necessary to modify the equation (2.7) into equation (2.8) as it was done in the work by Vorobyev et al. (2010). Here equation (2.7) is used to calculate diversity index.

DIV – not all industries are present in the city – according to (2.7)



DIV – not all industries are present in the city – according to (2.8)



Note: Gini index is defined as follows: Gini= $2-(3*x_1+2*x_2+x_3)$, $x_1<x_2<x_3$

HHI index is defined as follows: HHI= $x_1^2+x_2^2+x_3^2$

where x_j – a share of revenue of an industry *i* in a city

Figure 2.3. Comparison of diversity indices; numerical example of 3 industries: the 1st industry share is being changed, the 2nd and the 3rd ones have equal remaining shares

For the case when there is much less industries in the city than in the country, in this example it is assumed that there are 6 industries in the country. Then the value of index calculated using the equation (2.7) diminishes significantly, while the value of index (2.8) remains relatively higher. The reason is that in equation (2.7) comparison with industrial structure of the country is made. The same time, in equation (2.8) estimation is made using information on the city and keeping in mind that $div_t^z = \sum_{j=1}^{s}\left(\frac{1}{s}\right)^{\frac{1}{s}} = s^{1-\frac{1}{s}} \underset{s\to\infty}{\to} s$ (equation 2.9).

Below, DIV and HHI indices are compared, when the number of industries is being changed. In case of perfect equality both DIV and HHI reflect variety increasing as a number of industries *s* increases. However, if we turn to the opposite case when there is a dominant industry accounting for 90% of revenue then HHI fails to be very sensitive to the increase in the variety, while the index DIV performs better (Figure 2.4).

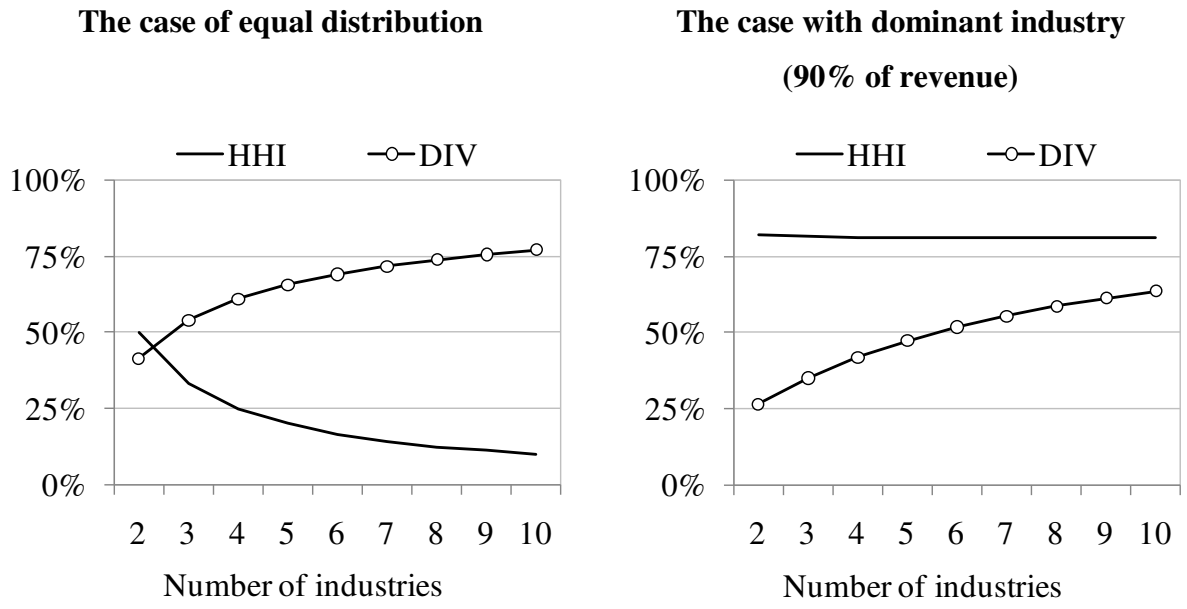


Figure 2.4. DIV (equation (2.8)) and HHI indices of diversity: a numerical example

We may conclude that HHI efficiently reflects diversity if the distribution is nearly uniform, but fails to provide good signals of growing variety if the distribution of industries is not uniform. Index DIV is sensitive to the variety in both cases. Table 2.16 below shows how enterprises in Russia are distributed according to localization and diversity faced by them.

Table 2.16. The number of firms located in cities with specific levels of diversity and localization

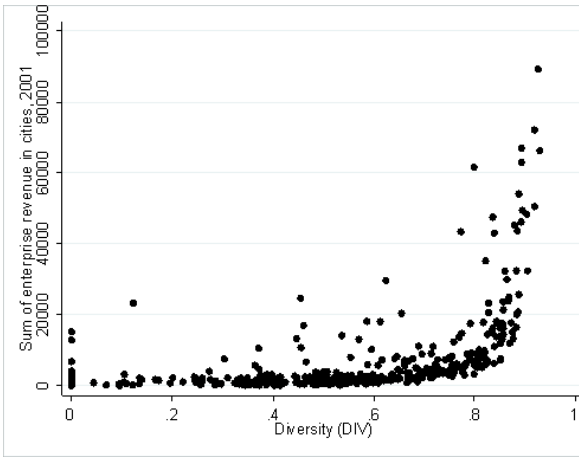
	Localization			
DIV	**< (mean-st.dev.)**	**mean +/-st.dev.**	**> (mean-st.dev.)**	**Total**
0-10%	**209**	**212**	**241**	**662**
10%-20%	9	11	3	23
20%-30%	**29**	10	-	39
30%-40%	**74**	11	9	94
40%-50%	**109**	34	5	148
50%-60%	**171**	38	10	219
60%-70%	**184**	47	19	250
70%-80%	**229**	113	3	345
80%-90%	**597**	**443**	**72**	**1112**
90%-100%	**180**	**562**	**518**	**1260**
total	**1791**	**1481**	**880**	**4152**

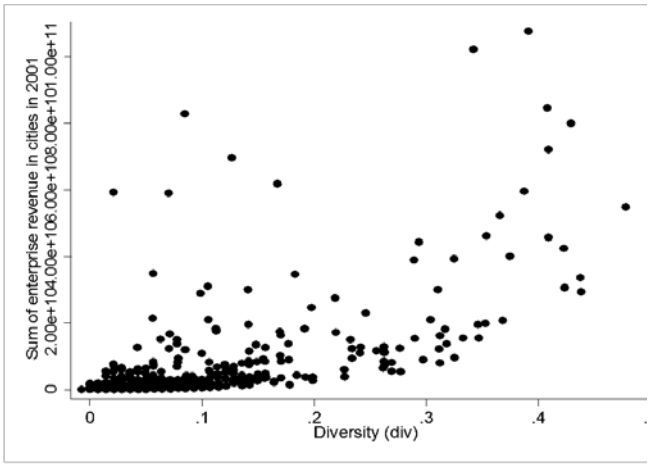
Source: Vorobyev et al. (2010). These results are received from database compiled by the Analytical agency 'Expert' and covering the years 2001-2005; diversity index is based on Equation (2.8)

Based on the firm level data for Russia, Vorobyev et al. (2010) found that at the lowest diversity levels[33] (0-10%) enterprises are distributed more or less evenly across cities with various localization levels. For the middle values of diversity there are more enterprises with a relatively low localization level. Finally, for the highest levels of diversity, distribution of enterprises becomes biased to the highest levels of localization. This pattern reflects the fact that the majority of enterprises in Russia are located in either monotowns or the largest cities.

Diversity is positively correlated with a city size. To show this, we plot diversity against the total revenue of enterprises in a city (see Figure 2.5). Larger city size means not only greater diversity but also greater localization (it is likely that in a larger city there are more enterprises of the same industry).

[33] In the research presented in these chapters, both diversity and localization reflect the situation in the city. However, in Vorobyev et al. (2010) the diversity index for the cities belonging to agglomerations, i.e. located within 60 km from the 'agglomeration centers' reflects the *overall diversity in the agglomeration.* The definition of agglomeration centers in Vorobyev et al. (2010) is the same as in this research: the cities with population exceeding million people and a number of other cities defined by experts as agglomeration centers.





Vorobyev et al. (2010)
Based on equation 2.8; *s* is number of industries in a city

Author's calculations
Based on equation 2.7; *s* is number of industries in a country

Figure 2.5. Diversity (DIV) vs. total enterprise revenue in cities, in 2001

Dynamics of economic activity concentration in Russia. All industries present in the sample are included into the analysis below (not only tradables; details about the sample are in Appendix B). Diversity index, as well as urbanization and localization indices, is based on revenue. However, it varies from 0 to 1, and its values do not depend on inflation, therefore dynamics of diversity over the years 1999 - 2008 can be compared. The sample allows to measure dynamics of diversity within manufacturing industries only.

The same way as Elison-Glaeser index considered above, Table 2.17 below shows that between 1999 and 2008 *diversity* initially increased, and then declined in all types of cities. Details on diversity levels in 17 agglomeration centers and comparison with the results received by Vorobyev et al. (2010) are in Table 2.18 below.

Table 2.17. Average diversity levels in different types of cities (for all industries in the database)

City type	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	Total
Agglomeration center	0.60	0.61	0.62	0.62	0.62	0.62	0.62	0.61	0.60	0.59	0.61
Within agglomeration; not center	0.09	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.09	0.09	0.10
Monotown within agglomeration	0.07	0.07	0.08	0.08	0.08	0.07	0.07	0.07	0.07	0.07	0.07
Monotown outside agglomeration	0.10	0.10	0.11	0.11	0.11	0.11	0.11	0.11	0.10	0.10	0.11
Other cities	0.15	0.16	0.16	0.17	0.17	0.17	0.17	0.16	0.16	0.15	0.16
Total	0.28	0.29	0.29	0.30	0.30	0.30	0.30	0.29	0.28	0.28	0.29

Table 2.18. Diversity index for the major cities*

City	1999	2004	2006	2008	2004**
Moscow	82%	83%	83%	81%	96%
St. Petersburg	69%	69%	67%	63%	96%
Samara	39%	42%	41%	39%	93%
Ekaterinburg	47%	50%	49%	48%	92%
Nizhny Novgorod	39%	44%	44%	39%	93%
Rostov-on-Don	36%	39%	38%	38%	90%
Novosibirsk	42%	44%	43%	42%	92%
Omsk	26%	32%	30%	30%	89%
Chelyabinsk	40%	39%	40%	37%	90%
Kazan’	42%	43%	42%	41%	90%
Krasnoyarsk	30%	33%	34%	32%	89%
Irkutsk	25%	28%	27%	25%	87%
Vladivostok	23%	25%	26%	25%	87%
Perm’	38%	39%	39%	39%	91%
Ufa	41%	48%	48%	48%	90%
Kaliningrad	33%	33%	33%	31%	89%
Vologda	15%	18%	18%	17%	82%

* Details about agglomeration centers are provided in Section 2.6 below. All industries present in the sample are included (not only tradables).

** From Vorobyev et al. (2010); diversity index is calculated based on equation (2.8), where *s* is the number of industries in the city. In this work diversity index is always calculated based on equation (2.7) where *s* is the number of industries in the country.

2.5. Home market potential

To construct the variable *Home market potential* (HMP), we consider the HPM in a territory itself and in all other territories of the country, weighed by distances. The first approach is to take into account potential demand *in a region* measured by gross regional product (GRP) (Table 2.20). In this case, to calculate HMP for a region, arising from the other regions, GRP of each of the other regions of the country is divided into the distance between the two regions’ capitals, and these expressions for each of the other regions are summed up. Then, together with GRP of the given region, this sum comprises the actual HMP faced by the enterprises located in the region. GRP is taken from the official website of the Russian Statistical Agency (Rosstat). The second approach is based on the potential demand *in a city*. It is measured with total revenue of firms in a city. As the measure of total revenue of firms in a city closely correlates with the diversity index included into the model, only firms’ revenue in the other cities is included into the regressions (Table 2.19).

In the tables below, there are means of home market potential, arising from the other cities, taking into account the geographical distances between cities. These means are calculated for different types of cities. The other part of the home market effect is revenue generated by all firms in the city itself; it is reflected in the urbanization index. The sample with data on city average wage contains observations for the years 2002-2008; the values of the city level home market potential are presented for these years.

Table 2.19. Home market potential arising from the other cities, mn roubles[34]

City type	2002	2003	2004	2005	2006	2007	2008
Agglomeration center	5,115.25	5,977.04	7,389.35	9,133.46	10,960.50	13,128.17	15,200.20
Within agglomeration; not center	10,744.84	12,821.71	15,695.92	18,840.77	22,687.14	28,193.68	32,897.68
Monotown within agglomeration	11,164.54	12,923.50	15,667.28	18,724.98	22,350.06	26,534.10	30,359.77
Monotown outside agglomeration	6,660.09	7,769.25	9,648.38	11,968.75	14,357.10	17,061.68	19,749.54
Other cities	6,140.80	7,219.77	8,850.32	10,831.02	12,792.17	15,396.32	17,924.29

As revenues are in nominal prices, it is impossible to understand the dynamics of home market potential. However, the table above shows distribution of home market potential on average for different types of cities.

The sample with regional level data used in Table 2.20 below, contains observations for the years 1999-2008. The part of the regional level home market potential is presented, generated by the other regions, and taking into account railroad distances between the regional capitals. It shows which regional HMP is faced by enterprises located in various types of cities, though this indicator is too general for city level analysis. As home market potential is based on nominal GRP, only conclusions about distribution of HMP on average for different types of cities can be made.

Table 2.20. Home market potential arising from the other regions, mn roubles[35]

City type	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Agglomeration center	2,726.72	4,089.89	5,126.82	6,223.11	7,629.11	9,841.57	12,557.21	15,564.86	19,210.14	23,612.96
Within agglomeration; not center	3,014.86	4,635.78	5,757.04	7,026.67	8,626.79	11,222.63	14,709.67	18,280.07	22,551.87	27,681.10

[34] Based on firms' revenue from SPARK-Interfax database; for all industries present in the database
[35] Based on gross regional product, Region level data, Rosstat

City type	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Monotown within agglomeration	3,418.67	5,315.56	6,598.35	8,011.33	9,860.54	12,871.66	16,925.90	20,936.16	25,605.12	31,276.43
Monotown outside agglomeration	3,067.76	4,683.67	5,812.01	7,092.35	8,711.55	11,302.29	14,769.78	18,461.37	22,862.96	28,071.51
Other cities	3,022.32	4,633.84	5,722.82	7,050.88	8,660.56	11,204.68	14,722.83	18,418.87	23,010.93	28,507.62

The patterns of home market potential distribution on average for different types of cities are similar in two tables above. However, on the city levels, differences between the types of cities are more pronounces. For example, it is evident that cities and towns, including monotowns, within agglomerations, have higher HMP than cities and towns outside agglomerations, because of distances to large cities and towns.

Table 2.21. Correlation matrix of agglomeration indices

	lloc	loc	core	lurbpq	urbpq	ldivpq	divpq	LQ	EGcities	EGind
lloc	1.0000									
loc	0.3200	1.0000								
core	-0.0180	0.0556	1.0000							
lurbpq	0.6312	0.3759	-0.4500	1.0000						
urbpq	0.4141	0.6143	-0.1114	0.6214	1.0000					
ldivpq	0.6415	0.3142	-0.4260	0.9119	0.5203	1.0000				
divpq	0.6277	0.4605	-0.2677	0.8593	0.7445	0.8788	1.0000			
LQ	-0.0516	-0.0149	0.1308	-0.0973	-0.0279	-0.0956	-0.0624	1.0000		
EGcities	-0.4687	-0.1823	0.3512	-0.5806	-0.3174	-0.7622	-0.6261	0.0700	1.0000	
EGind	0.0097	0.0923	-0.0811	0.1087	0.1219	0.1256	0.1576	0.0967	-0.0707	1.0000

From Table 2.21 it is clear that Diversity index is highly correlated with Elison-Glaeser index for cities; correlation is -0.62 with diversity index and -0.76 with logarithm of diversity index. As it was described above, diversity index is measured for cities; it ranges from 0 to 1, the value for the cities with the highest diversity striving to 1. The same time, Elison-Glaeser index ranges from zero, when the city is perfectly diverse, to two in the limit in case of high concentration. Therefore correlation is negative.

Localization quotient measures concentration of certain industry in a city; its correlation with Elison-Glaeser indices for cities and for industries is rather low (0.070 and 0.097). Urbanization index is measured for cities. As is would be expected, correlation of LQ with urbanization and with diversity is negative. However, it is also rather low, -0.097 with logarithm of urbanization and -0.0956 with logarithm of diversity. Localization index is measured for industry in a city. Correlation of localization index in logarithm with LQ is

negative and rather low (-0.05). Correlation of localization index with Elison-Glaeser index for industries is positive with the value of 0.09.

As for the indices used in the models in the following chapters, correlation of localization with urbanization and diversity is high (0.63 and 0.64 respectively; all indices are in logarithms). Correlation of urbanization and diversity is very high, 0.91 (both indices are measured in logarithms), i.e. in larger cities there is higher diversity of business activity, as it would be expected. Correlation of share of industry *j* in city *z* ('core') with localization is negative but rather low (-0.018). Its correlation with urbanization and diversity is negative and higher (-0.45 and -0.426), meaning that in cities with smaller scale and scope of economic activity the share of one industry in total firms' revenue is larger.

2.6. Territory classification

In Russia *a city or town* is a place with population 12 thousand people or more where no less than 85% of population is employed outside of agriculture. Based on Rosstat data for 1989-2007, 73% of Russian citizens live in cities and towns. By the end of 2010 in Russia there were 1100 cities and towns, and 1286 urban settlements, which have features both of small towns and of villages (Population census, Rosstat, 2010). The sample covers 1027 cities belonging to 76 out of 83 regions of the Russian Federation. The sample contains 17 agglomeration centers, 105 cities and towns belonging to agglomerations, and 94 monotowns. 24% of enterprises in the sample belong to agglomeration centers, and 29% are situated within agglomerations; 5% of enterprises are located in monotowns[36]. Settlements other than cities are not assigned codes in the sample; overall number of settlements is 1084.

In this work the following classification of the territory of Russia is used. Firstly, there are cities – agglomeration centers. Secondly, there are cities within agglomerations, which are not the agglomeration centers. Thirdly, there are monotowns located within agglomerations. Besides, there are monotowns located outside agglomerations. Finally, there are other cities, not included into the previous groups (cities and towns with population less than one million people, which do not belong to agglomerations and are not monotowns). Concerning the terms, 'cities' are relatively larger than 'towns', though here these words are used

[36] There are 7 federal districsts and 83 administrative units in Russia. Administrative division of the Russian Federation is organized as follows: 21 republics, 8 krais (territories), 47 regions (oblasts), 2 federal cities (Moscow and St. Petersburg), 1 autonomous region (oblast), 5 autonomous okrugs (areas), 1868 districts. Unlike federal districts, regions carry out the real fiscal, structural and social policy. There are 1100 cities (including federal cities), 329 town districts and okrugs (areas), 23160 rural administrations; 24.2 thousand municipal units (Rosstat, 1 January 2008, http://www.gks.ru/bgd/regl/B08_11/Main.htm)

interchangeably, for example, 'cities within agglomerations, not agglomeration centers' can be both relatively larger cities and relatively smaller towns.

We apply the definition that a city belongs to agglomeration if it is an agglomeration center or is located within 60 km distance from an agglomeration center. Cities within agglomerations, which are not the agglomeration centers, are cities belonging to agglomerations, i.e. a city located within 60 km from one of the agglomeration centers. Under agglomeration centers we mean the cities with population exceeding million people and a number of other cities defined by the experts as agglomeration centers; usually those are the cities with large economic activity. A number of agglomeration centers are defined based on classification by Gonchar (2008), who refers to the Ministry of Regional Development of the Russian Federation and the Center for Strategic Development North-West - presentation 'City agglomerations in Russia'. They are Moscow, St. Petersburg, Ekaterinburg, Rostov-on-Don, Nizhny Novgorod, Vologda, Irkutsk, Vladivostok, Novosibirsk, Krasnoyarsk. Gonchar (2008) adds Perm, Kaliningrad, Samara to this list based on her own judgment. In line with the criteria of population exceeding one million people we added to her list the following cities: Omsk, Kazan', Chelyabinsk, Ufa (overall, it gives 17 cities – agglomeration centers)[37].

Another types of cities are *monotowns*, those located within agglomerations and outside agglomerations. The definition and description of this type of cities is provided in chapter 1, section 1.3. Below maps of the Russian cities are presented; the borders of Russian regions are marked. Figure 2.6 presents the map of agglomeration centers.

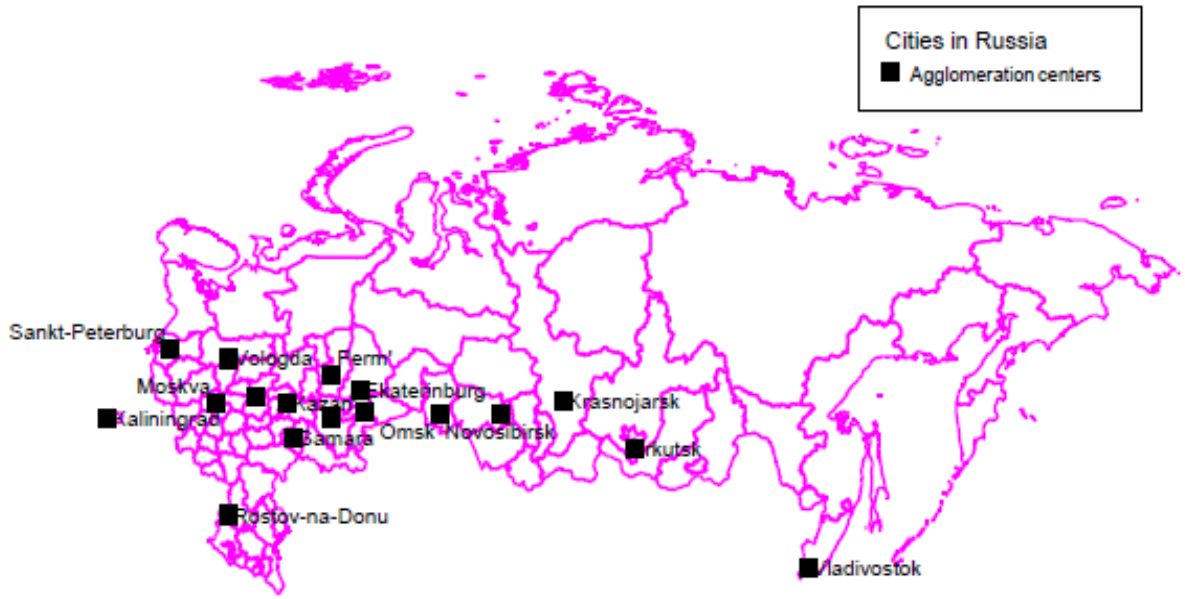


Figure 2.6. Agglomeration centers

[37] Rosstat, 1 January 2008, www.gks.ru – Russia in figures ('Rossiya v tsifrah')

On the map below (Figure 2.7), there are 1108 Russian cities, towns and urban settlements, including the agglomeration centers.

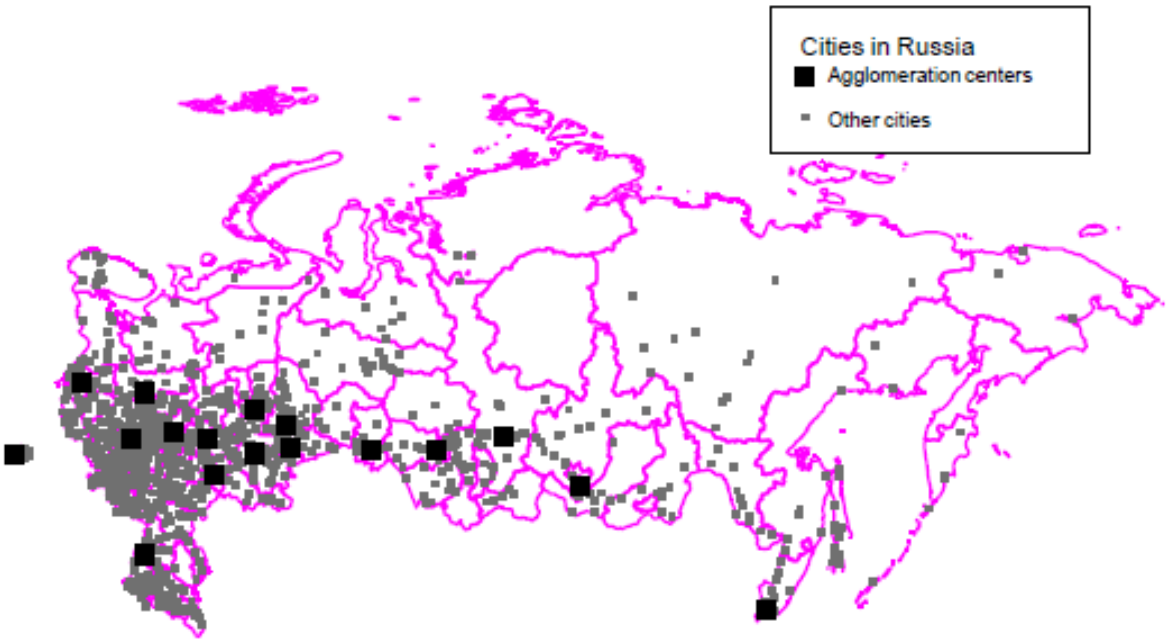


Figure 2.7. Agglomeration centers, and other cities and towns

The map below (Figure 2.8) illustrates location of railroads compared to the location of cities and towns in Russia.

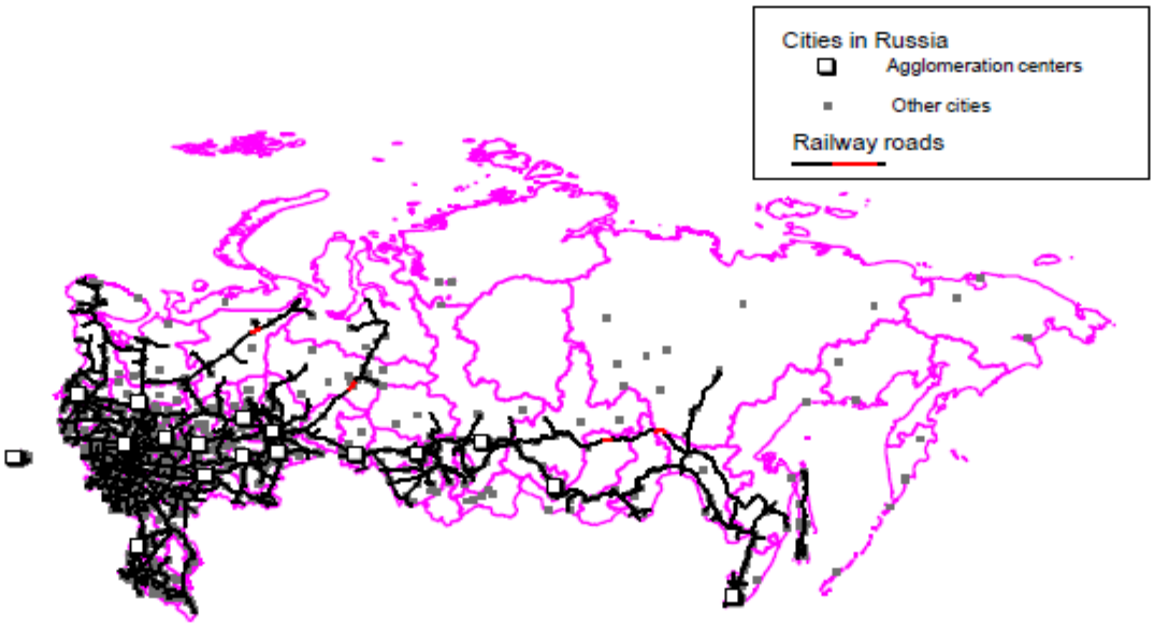


Figure 2.8. Railroads, and cities and towns

Below the map of monotowns and agglomeration centers is presented (Figure 2.9). The number of monotowns depends on the definitions of monotowns used by the experts: Zubarevich (2010) presents a list of 27 monotowns, while in the report *Monotowns in Russia* (2008) 146 monotowns of large business are presented.

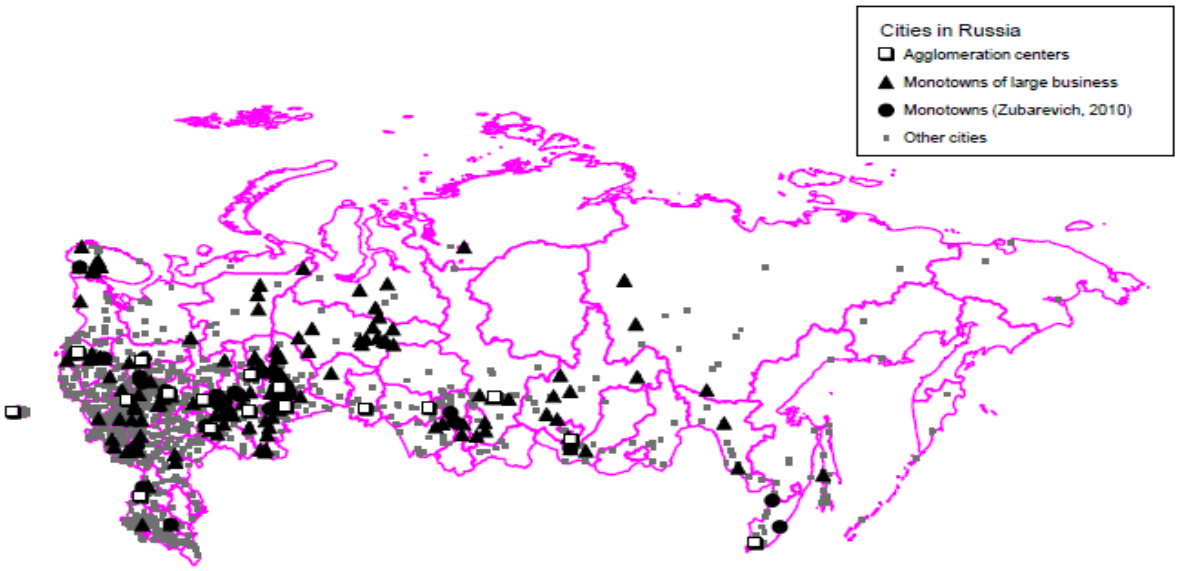


Figure 2.9. Agglomeration centers and monotowns

The majority of cities in Russia are located in the Western part of the country and along the Trans-Siberian railway; the monotowns are not an exception. The towns located in the Eastern part and outside of agglomerations, must be especially vulnerable to economic difficulties.

2.7. Estimation of the agglomeration economies

Agglomeration economies are estimated following three basic steps: finding an appropriate measure of economic performance; estimating the extent of agglomeration by the most appropriate measures; estimating the impact of agglomeration on economic performance through a regression of economic performance on agglomeration measures.

Performance measures. Agglomeration economies are often measured as productivity gains on a firm level. Generally, an empirical model can be based on a production function with a firm's output or value added on the left-hand side of a regression equation and agglomeration variables, production inputs and other control variables on the right-hand side. One of the possible specifications is given by Cobb-Douglas production

function with total factor productivity of the firm (TFP) affected by agglomeration externalities (Rosenthal, Strange, 2004), home market potential of a city, human capital level, transport infrastructure, business environment:

$$Y_{it} = A(E_{it})L_{it}^{\alpha}K_{it}^{\beta},$$

where Y_{it} is a firm's output at time *t*, L_{it} is labor, and K_{it} is capital, used by the firm *i* at time *t*; α is elasticity of output with respect to labor, and β is elasticity of output with respect to capital; E_{it} accounts for the externalities affecting the firm at time *t* and $A(E_{it})$ is total factor productivity of the firm (Vorobyev et al., 2010).

The factors enhancing productivity of enterprises are studied. In the models presented here, the sense is, which factors allow an enterprise to produce more, controlling for labour and capital employed by it. Vorobyev et al. (2010) provide an overview of performance measures. In a number of papers, enterprise revenue is a dependent variable, and production function has the log-log form (Henderson, 2003; Vakhitov and Bollinger, 2010). A more accurate indicator is value-added because it excludes expenditure on intermediary products. For instance, Martin et al. (2011) use value added in their research devoted to French enterprises. Employment growth is also used to measure performance. Besides, decisions of new establishments are analyzed in some works, i.e. birth of new plants and new establishment employment (Rosenthal and Strange, 2003). The value of marginal product of labor is another performance measure; it is equal to wages if perfect competition is assumed to be present on the labor market (Ciccone and Hall, 1996; Abel, 2010; Anastassova, 2006; Combes et al., 2008). Eriksson (2010) in his research on Swedish economy uses growth of labor productivity between the years 2001 and 2003 (labor productivity is defined as value added per employee). In this research logarithm of revenue is used as a dependent variable following Henderson (2003) and Vakhitov (2008).

2.8. Estimation issues

Endogeneity problem is the most difficult part of the estimation. It arises because explanatory variables may correlate with an error term. In this case OLS-estimates are biased. Endogeneity may occur for two major reasons (Martin et al., 2011).

Unobserved heterogeneity: there are some unobserved factors which affect both the output of enterprise and some explanatory variables. Because of this, the error term is correlated with explanatory variables, and explanatory variables capture part of the effect of the unobserved factors, leading to potential bias in coefficients estimated using the OLS model. For example, there could be unobserved or unmeasured city characteristics such as

climate, geographical position, resource endowment, transport infrastructure, policy of the municipal authority. They can affect both the performance of the local firms (dependent variable – output or revenue) and the attractiveness of a city for new firms (therefore affecting independent variables that reflect agglomeration of economic activity – localization and urbanization or diversity). It would cause bias in estimation of agglomeration variables coefficients. Another example is entrepreneur's characteristics that affect both choice of labour and capital, and choice of location with certain specific features related to productivity (Martin et al., 2011).

The use of firm level panel data allows addressing endogeneity problem more directly than industry level data. Particularly, one way to address unobserved heterogeneity is using individual fixed effects, which for the firms heterogeneity implies a set of firm dummy variables included into specification, or mean-differencing of (2.9). To measure industry, location and plant heterogeneity, Henderson (2003) includes industry-time and plant-location fixed effects. Another way to account for heterogeneity is to include into a regression additional variables measuring heterogeneity, for example, local characteristics.

In equation (2.9) the error term is assumed to be $u_{ti} = \varphi_i + \varepsilon_{ti}$, where φ_i are enterprise level fixed effects, and ε_{ti} is the error term which is assumed to be uncorrelated with independent variables, because φ_i captures enterprise invariant characteristics as well as characteristics of cities where they are located and industries to which they belong (enterprise do not change their location and industry over time):

$$\ln(revenue)_{ti}^{jz} = \beta_0 + \beta_1 \ln(\text{capital})_{ti}^{jz} + \beta_2 \ln(\text{labour})_{ti}^{jz} + \beta_3 core_t^{jz} + \beta_4 \left[core_t^{jz} \right]^2 + \beta_5 \ln\left(loc\right)_{ti}^{jz} +$$
$$+ \beta_6 \left[\ln\left(loc\right)_{ti}^{jz} \right]^2 + \beta_7 \ln\left(\text{urb}\right)_t^{jz} + \beta_8 \ln\left(hmp\right)_t^{z} + \beta_9 \ln\left(wage\right)_t^{z} + regional\_characteristics + \varphi_i + \varepsilon_{ti}, \quad (2.10)$$

Fixed effects model allows having endogenous regressors under condition that they are correlated only with a time-invariant component of the error (Cameron and Trivedi, 2009). Another way of addressing unobserved heterogeneity problem using panel data, with similar meaning as the fixed effect method described above, is time differencing (Martin et al., 2011). Then equation (2.9) would have the following form:

$$\triangle \ln(revenue)_{ti}^{jz} = \beta_0 + \beta_1 \triangle \ln(\text{capital})_{ti}^{jz} + \beta_2 \triangle \ln(\text{labour})_{ti}^{jz} + \beta_3 \triangle core_t^{jz} + \beta_4 \triangle \left[core_t^{jz} \right]^2 + \beta_5 \triangle \ln\left(loc\right)_{ti}^{jz} +$$
$$+ \beta_6 \triangle \left[\ln\left(loc\right)_{ti}^{jz} \right]^2 + \beta_7 \triangle \ln\left(\text{urb}\right)_t^{jz} + \beta_8 \triangle \ln\left(hmp\right)_t^{z} + \triangle \beta_9 \ln\left(wage\right)_t^{z} + \triangle regional\_characteristics + \triangle \varepsilon_{ti}, \quad (2.11)$$

Simultaneity. A possible source of simultaneity comes from an assumption that firms choose their location considering possible agglomeration advantages or disadvantages.

Successful businesses choose the cities where productivity is higher, and are therefore disproportionately found in agglomeration areas. As a result, the relationship between agglomeration and productivity will be biased upwards. The analogous conclusion can be made for weaker businesses, which would not survive in agglomeration, therefore they do not try to locate there, and their number is disproportionately lower than that of successful businesses (Vorobyev et al., 2010). Another potential source of simultaneity is productivity shocks that affect an enterprise (labour and capital that it uses) and the other enterprises in a city or region (localization and urbanization or diversity levels) (Martin et al., 2011).

A common way to deal with simultaneity problem is instrumental variables. For example, lagged values of explanatory variables and two-stage least squares method (2SLS) or generalized method of moments (GMM) can be used. Lagged values of explanatory variables can be applied if there is sufficiently large number of observation time-periods (Baum (2006), Cameron, Trivedi (2005)).

In the models presented in Chapters 3 and 4, both enterprise characteristics (labour and capital), and agglomeration indices may be correlated with error term. However, the assumption is made that national firms' location has a certain 'path dependence' on location decisions made in the times of planned economy. Therefore it is assumed that agglomeration variables reflecting concentration of national firms are exogenous. The part of home market potential that captures market size of the other cities is not likely to be endogenous. Transport infrastructure, particularly railroad infrastructure does not seem to be endogenous as well, because it was for the most part constructed during the times of planned economy for the reasons other than economic efficiency, and because it does not change rapidly with the change of economic activity allocation.

Equation (2.9) is estimated in Chapter 3, while equation (2.10) is estimated in Chapter 4, subdividing agglomeration externalities into those generated by national and foreign firms. In Chapter 3, endogeneity arising from unobserved firms' heterogeneity is addressed by using fixed effects. As enterprises do not change location and industry over time, these fixed effects capture specific features of industries and locations as well. It is assumed that endogeneity caused by simultaneity is not present in the model estimated in Chapter 3. In Chapter 4, first differencing is used to address unobserved heterogeneity, and first-differenced regressors are instrumented by their levels with 2 year lags using GMM procedure to deal with simultaneity. The econometric assumption behind this procedure is that idiosyncratic shock that occurred 2 years ago is orthogonal to the error terms e_{it}. Then the instruments are exogenous (Martin et al., 2011).

In Table 2.22 below the approaches to deal with endogeneity in data are presented.

Table 2.22. Estimation of agglomeration economies under endogeneity in data

	Source	Dependent variable	Independent variable	Estimation method	Instruments
1	Anastassova (2006)	wage	employment density	Two stage least squares (2SLS)	total land area of districts and population density in 1801
2	Henderson (2003)	output	agglomeration economies and inputs	Two stage least squares (2SLS) *and* Generalized method of moments (GMM) *proved to be ineffective* Industry-time and plant/location fixed effects	air quality attainment status of the territory, market potential of the territory, other cross-sectional MSA attributes
3	Henderson (1995)	own industry employment in 1987	own industry employment, concentration and diversity in 1970	Two stage least squares (2SLS)	variables not influenced by city-industry fixed effects
4	Martin et al. (2011)	value added	aglomeration economies and inputs	Instrumental variables method *and* Generalized method of moments (GMM) Time, employment area, industry and firm fixed effects	instrumental variables: lagged agglomeration variables

Source: Vorobyev et al. (2010)

2.9. Conclusion

In Chapter 2, the background for the following chapters was provided. Industries were subdivided into five groups based on the assumed differences in the agglomeration effects for them. Agglomeration indices were analyzed and descriptive statistics for some of them was calculated; diversity index applied in this work was presented. Cities were grouped into several types depending on specific conditions that enterprises face there: agglomeration centers, cities belonging to agglomeration, monotowns. The technique for estimation of agglomeration economies and potential issues of this estimation were described.

In Chapter 3, spatial concentration of economic activity is considered, and its impact on enterprise performance is studied.

3. The impact of spatial concentration on enterprise performance

3.1. Introduction

In Chapter 3, the objective is to reveal the impact of spatial concentration of business on enterprise productivity. Estimation is done by regressing the indicator of firm performance on a set of agglomeration indices, home market potential variable and control variables. The

following questions are considered. How does economic activity concentration affect economic performance? What is the role of localization, urbanization, diversity and home market potential? How do these effects vary across different types of cities, across industries and types of enterprises? Is theory confirmed by data? Which kinds of enterprises gain relatively more from agglomeration effects - national or foreign firms; private or public firms; firms created before or after the privatization, more productive or less productive firms? Do enterprises take into account agglomeration economies? Analysis of these questions will contribute to understanding, if economic policy can help, and if yes, how.

The following hypotheses are tested in this chapter. The first hypothesis is that both urbanization effects and home market potential positively affect enterprise performance. Variation of this impact across industries is also checked. The second hypothesis is that localization economies have an inverted U shape, i.e. start decreasing after some point due to congestion and excessive competition. The third hypothesis is that regional transport infrastructure and business climate affect enterprise performance. The fourth hypothesis is that human capital level is also fundamental for enterprise performance. To test these hypotheses, we analyze how firms' revenue is affected by agglomeration in city, home market potential, city average wages, transport infrastructure and business environment, controlling for the firms' inputs (labour and capital).

Concerning home market potential, it implies potential demand in the territory where the enterprise is located and in the other territories, taking into consideration that it is costly to reach the other territories. Home market potential is considered on the city level and is compared to the estimation based on regional level home market potential.

Literature review for this chapter is provided in Chapter 1. The chapter is structured as follows. In the next section, econometric model specification is provided. In Section 3.3, descriptive statistics is presented. Section 3.4 is devoted to the results of econometric estimation. Section 3.5 explores if localization level is optimal in Russia. Section 3.6 concludes.

3.2. Econometric model

The following model is estimated:

$$\begin{aligned}\ln(revenue)_{ti}^{jz} &= \beta_0 + \beta_1 \ln(\text{capital})_{ti}^{jz} + \beta_2 \ln(\text{labour})_{ti}^{jz} + \beta_3 core_t^{jz} + \beta_4 \left[core_t^{jz} \right]^2 + \\ &+ \beta_5 \ln\left(loc\right)_{ti}^{jz} + \beta_6 \left[\ln\left(loc\right)_{ti}^{jz} \right]^2 + \beta_7 \ln\left(\text{urb}\right)_t^{jz} + \beta_8 \ln\left(hmp\right)_t^{z} + \beta_9 \ln\left(wage\right)_t^{z} + \\ &+ \beta_{10} \ln\left(a\_road\right)_t^{r} + \beta_{11} \ln\left(rw\_road\right)_t^{r} + \beta_{12} \ln\left(\text{busnenvrisk}\right)_t^{r} + \varphi_i + \varepsilon_{ti},\end{aligned} \tag{3.1}$$

where *j* is industry index, *z* is city index, *r* is regional index, and *i* is firm index. Panel regression with enterprise fixed effects is applied. Estimation is done by regressing the logarithm of firm's revenue on the firm's characteristics, market size, and agglomeration indices. Specific features of industries and territories are reflected in the enterprise fixed effects, as in the sample, firms do not change location or industry. Hausman specification test showed that the fixed effects model is preferable over random effects method of panel data estimation. Cobb-Douglas function proved to reflect data better than Translog function. Robust standard errors were used to deal with the heteroscedasticity problem, as normality and homoscedasticity of the random error terms were not assumed. The models with industry level, regional level, and time fixed effects were also tested. Enterprise fixed effects were found to be the most relevant here, and are used in the models discussed below.

Table 3.1. Definition of variables used in the econometric analysis

Variable	Definition
Enterprise level characteristics.	
$\ln(revenue)_{ti}^{jz}$	Logarithm of enterprise revenue (revenue is in roubles)
$\ln(\text{capital})_{ti}^{jz}$	Logarithm of fixed assets (fixed assets are in roubles)
$\ln(\text{labour})_{ti}^{jz}$	Logarithm of labour force (number of employees)
Agglomeration indices[38]. City level.	
$\ln\left(loc\right)_{ti}^{jz}$	Localization coefficient (Equation (2.1))
$core_t^{jz}$	A share of an industry *j* in the total revenue in a city *z*
LQ_{jz}	Localization quotient (Equation (2.2))
$\ln\left(div\right)_t^z$	Diversity coefficient (Equation (2.7))
$\ln\left(\text{urb}\right)_t^{jz}$	Coefficient of urbanization (Equation (2.3))
Home market potential. Regional level: HMPregion= $\ln\left(grp\right)_t^r + \ln\left(hmp\right)_t^r$.	
$\ln\left(grp\right)_t^r$	Logarithm of gross regional product (GRP is in mn roubles)
$\ln\left(hmp\right)_t^r$	The sum of gross regional products of other regions divided into the distances between the region *R* and the other regions. $hmp_t^R = \sum_{r \neq R} \frac{grp_t^r}{dist_{R,r}}$, where $dist_{R,r}$ are the *railroad* distances between the regional capitals

[38] All coefficients ('core', localization quotient, localization index, diversity and urbanization indices) are measured based on 3-digit level of OKVED classification

Variable	Definition
Home market potential. City level: HMPcity= $\ln\left(div\right)_t^z + \ln\left(hmpcity\right)_t^z$.	
$\ln\left(hmpcity\right)_t^z$	Total revenue in all industries in all cities other than the city where the firm is located, divided into the distances between the city *Z* and the other cities: $hmpcity_t^Z = \sum_{z \neq Z} \frac{revenue_t^z}{dist_{Z,z}}$, where $dist_{Z,z}$ are the *physical*_distances between the cities
City characteristics.	
$\ln\left(\text{wagecity}\right)_t^z$	Average monthly nominal wage, payroll, roubles (proxy for human capital)
Regional transport infrastructure.	
$\ln\left(\text{a\_road}\right)_t^r$	Density of auto roads in region r, end of year, km of roads per 1000 sq km of territory
$\ln\left(rw\_\text{road}\right)_t^r$	Density of railroads in region r, end of year, km of roads per 1000 sq km of territory
Regional investment climate (Indices by analytical agency 'Expert')[39].	
$\ln\left(\text{expertinvptl}\right)_t^r$	Investment potential in region r at time t (Regional investment potential by the Analytical agency Expert)
$\ln(\text{busnenvrisk})_t^r$	Business environment risks, region r, time t (Regional investment risk by the Analytical agency Expert)
Other regional characteristics.	
$\ln\left(\text{wageR}\right)_t^r$	Average monthly nominal wage of organizations' employees, payroll

Localization loc_{jz} in a quadratic form is included into regression to test the inverted U shape of agglomeration economies suggested in the literature. This form seems to be more relevant for localization economies than for the urbanization ones, particularly due to competition among the firms of the same industry for labour force and for the other resources.

To measure diversity, the diversity index introduced in this paper (Equation (2.7)) is used. It is calculated for a city based on 3-digit industrial classification (OKVED). $\ln\left(div\right)_t^z$ is used interchangeably with $\ln\left(\text{urb}\right)_t^{jz}$, as cities with higher urbanization tend to be more diversified; correlation between these coefficients is about 90%.

To analyze the issues connected with the dominating role of an industry in a city we use a variable $core_{jz} = \frac{revenue_{jz}}{revenue_z}$, a share of an industry *j* in the total revenue in a city *z*. This index also represents the numerator in the localization quotient.

The variable accounting for regional business environment risks *ln(busnenvrisk)* is included into the model, because the higher are the 'risks' in the region, the more time

[39] Details are in Appendix C 'Data sources'

entrepreneurs and managers spend to deal with these risks, and the less time is devoted to the production process. Moreover, higher risks divert attention from all kinds of innovations, such as improvement of products, technologies, and management processes.

The variable accounting for investment potential *ln(expertinvptl)* proved to be insignificant, and R squared (between and overall) is lower with this variable, so it is excluded from the model. Standard deviation of investment risk is 1.11 and of investment potential is 1.31 (both indices vary from 0 to 4.49), i.e. investment potential varies across regions. However, according to the Analytical agency 'Expert', investment potential became important for investors only starting from 2008, when investment risks were considerably lowered as a result of regional government policy during previous years. Besides, the federal center started to carry out more measures aimed at regulating investment climate in the regions (such as creating 22 special economic zones in 20 regions; enlarged list of 42 strategic industries were foreign investors access is limited by the Federal law №53-FZ, 29.04.2008). Due to concentration of legislative initiatives in the federal center, differences among regions diminished, one of the consequences being that legislative risk always considered to be significant by investors became relatively less important in 2008 (Expert, 2008). As the sample under consideration contains data up to 2008, business risks are relevant. Regional and city wages (interchangeably) were included as a proxy for human capital.

3.3. Data and descriptive statistics

In Tables 3.2 and 3.3 below descriptive statistics is presented. The sample in Table 3.2 contains enterprises belonging to tradable industries, for which city level data is available. Those are the cities with population over 100,000 people; data covers the years 2002-2008.

Table 3.2. Descriptive statistics for tradable industries; firms, for which *city* level data is present

Variable	Obs	Mean	Std.Dev.	Min	Max
revenue	56705	5.66e+08	3.54e+09	1,000.00	1.83e+11
fixed_assets	55479	1.40e+08	9.78e+08	23.00	5.49e+10
labour	53360	526.23	1,651.34	1.00	81,000.00
lrevenue	56705	17.86	2.26	6.91	25.94
lfixed_ass~s	55479	16.03	2.58	3.14	24.73
llabour	53360	5.11	1.50	0.00	11.30
core	60113	0.04	0.10	0.00	1.00

Variable	Obs	Mean	Std.Dev.	Min	Max
core2	60113	0.01	0.07	0.00	1.00
lloc	54542	16.53	8.73	0.00	26.56
lloc2	54542	349.64	208.91	0.00	705.66
ldivpq	60113	-1.05	0.71	-4.96	-0.18
divpq	60113	0.43	0.25	0.01	0.84
lurbpq	60113	25.11	1.94	12.58	28.38
lhmecity	60113	23.02	0.53	21.05	24.57
lwagecity	60113	9.15	0.57	7.46	12.05
la_road	60050	5.00	0.96	1.70	6.28
lrw_road	59933	5.37	0.90	2.20	6.35
lbusnenvrisk	60113	2.77	1.13	0.00	4.47

In Table 3.3 below there is descriptive statistics for the sample that contains enterprises belonging to tradable industries, for which region level data is available. Data covers the years 1999-2008.

Table 3.3. Descriptive statistics for tradable industries; firms, for which *region* level data is present

Variable	Obs	Mean	Std.Dev.	Min	Max
revenue	119309	4.37e+08	3.50e+09	57.00	3.12e+11
fixed_assets	116956	1.34e+08	2.25e+09	16.00	6.02e+11
labour	112855	490.93	1,585.11	1.00	84,918.00
lrevenue	119309	17.37	2.36	4.04	26.47
lfixed_assets	116956	15.76	2.61	2.77	27.12
llabour	112855	5.00	1.55	0.00	11.35
core	127622	0.08	0.17	0.00	1.00
core2	127622	0.04	0.14	0.00	1.00
lloc	113864	12.78	9.93	0.00	26.62
lloc2	113864	261.80	222.25	0.00	708.62
ldivpq	127622	-1.66	1.16	-4.99	-0.18
divpq	127622	0.32	0.27	0.01	0.84
lgrp	127622	12.29	1.48	7.40	15.95
lhmeR	118360	9.08	0.83	5.93	11.14
lwage	127622	8.60	0.83	6.22	10.55
la_road	127622	4.82	1.07	-0.22	6.28
lrw_road	126916	5.24	1.01	-0.69	6.37
lbusnenvrisk	127622	2.85	1.17	0.00	4.48

As in the cities with population over 100,000 people, business activity is larger and is probably more diversified than on average in the regions. Urbanization, diversity and localization levels are also larger on average, as it is seen from Tables 3.2 and 3.3 above. A share of an industry *j* in the total revenue in a city *z* is smaller in these cities. Average wage is larger, as it would be expected.

Correlations of the variables for the samples analyzed in Section 3.3 are presented in Appendix A. The results in Table A.1 in Appendix A (for the sample of cities with population over 100,000 and for all cities) show that in the larger cities, i.e. the cities where urbanization is higher, localization and diversity levels are also higher. Correlation between urbanization and localization is 0.543, and between urbanization and diversity 0.895. The share of an industry *j* in the total revenue in a city *z* is negatively correlated with urbanization: -0.1975. Wages are also higher in the larger cities, correlation between urbanization and wages being 0.534.

The results in Tables A.1 and A.2, i.e. for the sample of cities with population over 100,000 and for all cities, show that wage is positively correlated with diversity (0.2712 and 0.18 in Tables A.1 and A.2 respectively). Diversity is negatively correlated with regional business environment risks (-0.4175 and -0.4047 in Tables A.1 and A.2 respectively). Regional business environment risks are negatively correlated with city size (urbanization), -0.4682, and with gross regional product, -0.4275. Automobile and railroad density are positively correlated both with urbanization level and with gross regional product. These correlations are in line with economic intuition.

3.4. Estimation results

Equation (3.1) was estimated using panel data model with enterprise fixed effects. Fixed effects are applied to deal with endogeneity arising from the unobserved heterogeneity. As it will be commented in section 3.5 below, endogeneity caused by simultaneity should not arise in the cases considered in this chapter. The results of econometric testing of the equation (3.1) are presented in Tables 3.4 - 3.14. In the analysis below we compare the models with home market potential (HMP) based on city data and based on regional data. On one hand, we are interested more in the city level due to large size of the Russian regions. On the other hand, city level HMP is calculated using geographical distances, while the regional level one is calculated based on railroads distances between regional capitals, this way reflecting the actual transportation possibilities.

Table 3.4. Regression results with HMP based on city level data[40]
Dependent variable: ln(*revenue*)

	(1) All tradables 1	(2) All tradables 2
lfixed_assets	0.236***	0.234***
	(22.68)	(22.59)
llabour	0.431***	0.428***
	(32.07)	(32.00)
core	8.738***	9.574***
	(14.30)	(15.31)
core2	-6.258***	-6.628***
	(-7.79)	(-8.20)
lloc	0.0287***	0.0312***
	(4.01)	(4.42)
lloc2	-0.00106***	-0.00126***
	(-2.77)	(-3.37)
ldivpq	0.565***	
	(5.15)	
lurbpq		0.282***
		(10.69)
lhmecity	0.585***	0.451***
	(11.53)	(8.44)
lwagecity	0.118***	-0.000303
	(2.86)	(-0.01)
la_road	0.0652**	0.0445*
	(2.39)	(1.66)
lrw_road	-0.500	-0.466
	(-1.57)	(-1.54)
lbusnenvrisk	-0.0166***	-0.00805
	(-3.12)	(-1.53)
_cons	-0.0152	-3.588*
	(-0.01)	(-1.90)
firm fixed effects	yes	yes
N	49779	49779
adj. R^2	0.425	0.428
sigma_u	1.172	1.153
sigma_e	0.523	0.522
rho	0.834	0.830

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

[40] Total revenue of firms in a city

The results in Table 3.4 above show that localization, urbanization, diversity (i.e. agglomeration level in a city) and home market potential arising from the other cities are significant. A share of an industry *j* in the total revenue in a city *z* ('core' coefficient) and localization coefficient have an inverted U shape. 1% increase in 'core' leads to 0.087-0.096% increase in firm's revenue in Models (1) – (2). But when 'core' reaches 8.738/(2*6.258)=0.698 (for Model (1)), further 1% increase in 'core' leads to decrease in revenue by 0.063% compared to the optimum level of 'core', i.e. the 'core' effect stays positive but decreases. For Models (2) the threshold is 0.72.

As for localization coefficient, its increase by 1% leads to increase in firm's revenue by around 0.029-0.03%. In case of Model (1), when the level of localization 0.0287/(2*(0.00106))=13.538 is reached i.e. a very low level of total revenue in industry in city equal to exp(13.538)=757,467 roubles, then each 1% further increase in localization starts decreasing firm's revenue by 0.00065%. In other words, the effect decreases although stays positive until 0.0287/(0.00106)=27.075, i.e. exp(27.075)=0.57 bn roubles[41]. In Model (2) the optimum localization level is 12.381.

Economic activity in industries other than that where a firm works is reflected by diversity and urbanization interchangeably. The values of these coefficients are robust to the changes in specification. Diversity is significant and positive, showing that 1% increase in diversity leads to 0.565% increase in enterprise revenue. It means that if we compare a 'Center' with diversity approaching to 1 and a 'Periphery' with diversity approaching to 0, i.e. were diversity would be twice lower, the difference in productivity would approach 56.5%. The results are robust when Moscow is excluded.

Urbanization is significant and positive too, 1% increase in urbanization leading to 0.28% increase in enterprise revenue. The results for urbanization *economies* are consistent with the findings discussed in Chapter 1 (literature review). Enterprises located in the 'Centre' have an increase in productivity vs. 'periphery' by 20-50% (Okubo, Tomiura, 2010). For Russia, based on data for 2005-2006, it was found that doubling city size increases firms' productivity by 5% (Enterprises and markets, 2010). Firm located within an agglomeration vs. outside an agglomeration has an increase in productivity by 46% (Gonchar, 2010). With increase in diversity in a city by 0.1 (diversity varies between 0 and 1) the firm revenue growth rate increases by 4.7% over 3 years (2001 to 2004) and by 1.5% annually (Vorobyev et al., 2010). Rosenthal and Strange (2004) summarize that doubling city size (measured for USA as population of a Standard Metropolitan Statistical Area) leads to 3 – 8% increase in firm's productivity.

[41] Billion according to the long scale

HMP on city level consists of total revenue of firms in a city, excluding the revenue of the firm itself (urbanization coefficient) and the total revenue of firms in the other cities, the distances between cities being taken into account (variable 'lhmecity'). As for HMP arising from the other cities, its increase by 1% leads to increase in firm's revenue by 0.451 – 0.585% in specifications (1) – (2). The effect of wage is positive in specification with diversity index, possibly implying that firms benefit from human capital quality reflected in higher wages. 1% increase in city average wage leads to 0.118% increase in the firm's revenue. The effect of the automobile road density is positive and significant; 1% increase in automobile road density leads to 0.044-0.065% increase in the firm's revenue. The effect of the railroad density turned out to be insignificant or negative. In specification with diversity coefficient, business environment risk index is significant on 99% confidence interval showing that 1% increase in the regional business environment risk leads to decrease in the firm's revenue by 0.017%. In the specification below with city level HMP, urbanization index is chosen over diversity index because it is more suitable to reflect HMP within a city.

In Table 3.5 below HMP on the regional level is presented. As it was mentioned above, in this case railroad distances between regional capitals are used to calculate HMP arising from the other regions, while in case of city level HMP, geographical distances are used. Therefore it will be useful to compare the results.

Table 3.5. Regression results with HMP based on regional data[42]
Dependent variable: ln(*revenue*)

	(1) All tradables 1	(2) All tradables 2
lfixed_assets	0.200*** (28.02)	0.197*** (27.72)
llabour	0.516*** (43.78)	0.509*** (43.40)
core	6.757*** (27.06)	7.608*** (30.47)
core2	-4.503*** (-17.33)	-4.722*** (-18.86)
lloc	0.0202*** (3.70)	0.0211*** (3.91)
lloc2	-0.000652** (-2.33)	-0.000946*** (-3.38)

[42] Gross regional product (Rosstat)

	(1) All tradables 1	(2) All tradables 2
ldivpq	0.482*** (13.67)	
lurbpq		0.270*** (21.20)
lgrp	0.200*** (5.76)	0.0695** (2.01)
lhmeR	0.154*** (4.20)	0.0476 (1.32)
lwageR	0.306*** (10.83)	0.325*** (11.89)
la_road	-0.0431* (-1.76)	-0.0689*** (-2.88)
lrw_road	-0.171** (-2.40)	-0.164** (-2.33)
lbusnenvrisk	-0.0137*** (-2.94)	-0.0137*** (-2.96)
_cons	6.632*** (16.30)	2.082*** (4.62)
firm fixed effects	yes	yes
N	96869	96869
adj. R^2	0.561	0.568
sigma_u	1.069	1.054
sigma_e	0.575	0.571
rho	0.776	0.773

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The results in Table 3.5 above show that agglomeration effects are significant in specification both with diversity and urbanization, and HMP arising from the other regions is significant in specification with diversity. A share of an industry *j* in the total revenue in a city *z* ('core' coefficient) and localization coefficient have an inverted U shape, as they did in specification with HMP on city level (Table 3.4). 1% increase in 'core' leads to 0.067-0.076% increase in firm's revenue in Models (1) – (2), but when 'core' reaches 6.757/(2*4.503)=0.75 (for Model (1)), further 1% increase in 'core' leads to decrease in positive effects by 0.045%. For Models (2) the threshold is 0.805.

As for localization coefficient, its increase by 1% leads to increase in firm's revenue by around 0.02% in both specifications. In case of Model (1), when the localization level 0.0202/(2*(0.000652))=15.491 is reached, i.e. exp(15.491)=5.3 mn roubles, 1% further

increase in localization starts decreasing firm's revenue by 0.000652%, though the effect stays positive until 0.0202/(2*(0.000652))=30.98, which means exp(30.98)=28.5 bn, exceeding the result received with city level HMP. In Model (2) the threshold is 11.15. 'Core' and localization coefficients are slightly lower than in specification with city level HMP.

Economic activity in industries other than that where a firm works is reflected by diversity and urbanization indices interchangeably. Diversity is significant and positive, showing that 1% increase in diversity leads to 0.482% increase in enterprise revenue. Urbanization is significant and positive too, 1% increase in urbanization leading to 0.27% increase in enterprise revenue. HMP generated by the region itself is reflected by gross regional product. The results show that 1% increase in GRP leads to 0.0695 – 0.2% increase in the firm's revenue. As for HMP arising from the other regions, it is significant in specification with diversity coefficient (Model (1)) on 99% confidence interval. Its increase by 1% leads to increase in firm's revenue by 0.154%, which is lower than in specification with city level HMP considered above.

The effect of wage is either positive or insignificant, possibly implying that firms benefit from human capital quality reflected in higher wages. The results show that 1% increase in regional average wage leads to 0.306-0.325% increase in the firm's revenue, compared to 0.118% increase in the models with *city level* HMP and wage (Table 3.4). The effect of the automobile road density changed the sign compared to the models in Table 3.4, where it was positive and significant. The effect of the railroad density turned out to be negative as it was in Table 3.4, and in this case always significant. Business environment risk again has negative and significant impact on firm's revenue; in this case it is always significant, on 99% confidence interval. 1% increase in business environment risk in a region leads to decrease in the firm's revenue by 0.0137%, which is slightly lower than in Table 3.4.

Table 3.6. Regression results with HMP based on city level data, by class of industry
Dependent variable: ln(*revenue*)

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
lfixed_assets	0.234***	0.202***	0.233***	0.220***	0.235***	0.223***
	(22.59)	(19.01)	(11.14)	(6.14)	(2.69)	(14.76)
llabour	0.428***	0.502***	0.407***	0.412***	0.513***	0.439***
	(32.00)	(34.98)	(14.94)	(12.36)	(4.80)	(21.90)
core	9.574***	23.37***	12.04***	13.58***	10.38***	8.474***
	(15.31)	(13.76)	(8.91)	(5.35)	(5.17)	(9.93)

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
core2	-6.628***	-34.99***	-9.615***	-13.81***	-3.245*	-9.140***
	(-8.20)	(-9.18)	(-6.93)	(-3.02)	(-1.68)	(-7.94)
lloc	0.0312***	0.0220	0.0691***	0.0453**	0.182***	0.00447
	(4.42)	(1.47)	(5.18)	(2.40)	(5.04)	(0.27)
lloc2	-0.00126***	-0.00116	-0.00335***	-0.00230**	-0.0106***	-0.000223
	(-3.37)	(-1.41)	(-4.94)	(-2.33)	(-5.32)	(-0.27)
lurbpq	0.282***	0.159***	0.301***	0.326***	0.420*	0.306***
	(10.69)	(3.10)	(5.63)	(4.74)	(1.78)	(6.38)
lhme city	0.451***	0.322***	0.629***	0.463***	-1.165	0.292***
	(8.44)	(3.11)	(5.99)	(3.68)	(-1.22)	(2.89)
lwage city	-0.000303	0.196**	-0.139*	-0.0315	1.080*	0.175**
	(-0.01)	(2.47)	(-1.77)	(-0.34)	(1.76)	(2.28)
la_road	0.0445*	0.0236	0.00918	0.0307	-0.00560	0.0992**
	(1.66)	(0.39)	(0.19)	(0.54)	(-0.03)	(1.97)
lrw_ road	-0.466	-0.238	-0.579	1.738**	2.942	-0.998*
	(-1.54)	(-0.47)	(-1.24)	(2.02)	(0.94)	(-1.90)
lbusn envrisk	-0.00805	-0.00794	-0.0202*	-0.000430	-0.0806	-0.0249***
	(-1.53)	(-0.68)	(-1.79)	(-0.04)	(-1.59)	(-3.16)
_cons	-3.588*	-0.450	-5.810*	-15.76***	2.818	0.447
	(-1.90)	(-0.13)	(-1.89)	(-3.01)	(0.12)	(0.13)
firm fixed effects	yes	yes	yes	yes	yes	yes
N	49779	5590	12708	6741	746	17078
adj. R^2	0.428	0.331	0.464	0.453	0.448	0.442
sigma_ u	1.153	1.018	1.289	1.849	2.595	1.237
sigma_ e	0.522	0.478	0.498	0.497	0.608	0.529
rho	0.830	0.819	0.870	0.933	0.948	0.845

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Concerning the groups of industries (Table 3.6 above), share of an industry *j* in the total revenue in a city *z* ('core' coefficient) turned out to be significant for all industries. It has an inverted U shape and is relatively more important for traditional industries, followed by basic machinery industries, basic materials, integrating machinery and science based

industries. Localization coefficient has an inverted U shape; it is the largest for integrating machinery, followed by basic materials and basic machinery. Both these coefficients reflect the impact of city's specialization on a certain industry on the performance of an enterprise working in this industry.

Urbanization coefficient is significant for all industries. It is the most important for integrating machinery industries and basic machinery, followed by science based, basic materials and traditional industries. Home market potential arising from the other cities is significant for all industries except for integrating machinery. However, for this industry there are not so many observations (746), therefore the results might be inaccurate. The impact of HMP is the highest for basic materials industries, followed by basic machinery, traditional and science based industries. Overall, HMP in the city and in the other cities turned out to be the most important for basic materials, basic machinery, followed by science based and traditional industries. It may be partly connected to costs of transporting goods, differing among industries.

Wage effect turned out to be positive and significant for integrating machinery, traditional industries, and science based industries, and negative for basic materials industries. On the industry group level, the effect of automobile road density is positive for science based industries; the effect of railroad density is negative for science based industries, but turned out to be positive and significant for basic machinery. Regional business environment risk is found to produce negative effect on science based and basic materials industries, for other industries it is insignificant.

Table 3.7. Regression results with HMP based on regional data, by class of industry
Dependent variable: ln(*revenue*)

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
lfixed_assets	0.200***	0.162***	0.201***	0.183***	0.173**	0.209***
	(28.02)	(23.67)	(15.76)	(7.82)	(2.19)	(15.80)
llabour	0.516***	0.572***	0.489***	0.522***	0.728***	0.477***
	(43.78)	(54.04)	(21.99)	(16.76)	(6.99)	(23.25)
core	6.757***	7.526***	7.509***	7.105***	4.095***	6.161***
	(27.06)	(24.78)	(15.68)	(10.12)	(2.60)	(9.48)
core2	-4.503***	-5.258***	-5.207***	-4.698***	-2.164	-4.165***
	(-17.33)	(-15.06)	(-10.23)	(-6.89)	(-1.51)	(-3.51)
lloc	0.0202***	0.0295***	0.0628***	0.0652***	0.104***	0.0142*
	(3.70)	(2.90)	(6.71)	(4.81)	(3.91)	(1.66)

lloc2	-0.000652**	-0.00143***	-0.00310***	-0.00358***	-0.00592***	
	(-2.33)	(-2.68)	(-6.24)	(-4.99)	(-4.11)	
ldivpq	0.482***	0.361***	0.558***	0.531***	0.658**	0.639***
	(13.67)	(6.40)	(8.56)	(5.57)	(2.38)	(5.77)
lgrp	0.200***	-0.0203	0.180***	0.240**	-0.0648	0.108
	(5.76)	(-0.36)	(2.88)	(2.48)	(-0.27)	(1.55)
lhmeR	0.154***	0.212***	0.438***	0.109	-0.366	0.115
	(4.20)	(3.44)	(6.38)	(1.07)	(-1.36)	(1.60)
lwageR	0.306***	0.435***	0.0713	0.385***	1.168***	0.506***
	(10.83)	(7.20)	(1.44)	(5.00)	(6.33)	(8.58)
la_road	-0.0431*	0.00462	-0.0647	-0.0771	-0.131	-0.0149
	(-1.76)	(0.09)	(-1.54)	(-1.30)	(-0.76)	(-0.27)
lrw_ road	-0.171**	-0.183***	-0.371**	0.590	-0.921	-0.154
	(-2.40)	(-2.66)	(-2.15)	(1.03)	(-0.68)	(-1.35)
lbusn envrisk	-0.0137***	-0.0130	-0.0270***	-0.0140	-0.0801**	-0.0254***
	(-2.94)	(-1.51)	(-2.87)	(-1.14)	(-2.49)	(-3.17)
_cons	6.632***	7.455***	7.865***	2.795	11.69*	5.802***
	(16.30)	(16.57)	(8.41)	(0.93)	(1.75)	(8.08)
firm fixed effects	yes	yes	yes	yes	yes	yes
N	96869	12884	24992	12223	1601	27109
adj. R^2	0.561	0.473	0.608	0.576	0.585	0.577
sigma_ u	1.069	0.955	1.101	1.018	1.373	1.123
sigma_e	0.575	0.524	0.523	0.558	0.655	0.618
rho	0.776	0.768	0.816	0.769	0.815	0.768

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Concerning the groups of industries (Table 3.7 above), the results are similar to those in Table 3.6, for estimation with city level HMP and city wage. Share of an industry *j* in the total revenue in a city *z* ('core' coefficient) is significant for all industries. It has an inverted U shape and is relatively more important for traditional industries, followed by basic materials and basic machinery industries; then come science based and integrating machinery industries. Localization coefficient has an inverted U shape for all industries except for science based industries, for which it is positive and significant on 90% confidence interval. It is the largest for integrating machinery, followed by basic machinery, basic materials, and traditional industries. For traditional and science based industries, the coefficient was not

significant in Table 3.6. Both these coefficients reflect the impact of city's specialization on a certain industry on the performance of enterprise working in this industry.

HMP used in this estimation is measured as gross regional product (GRP) of the region where an enterprise is situated and GRP of the other regions, taking into account railroad distances between the regions' capitals. GRP is positive and significant for basic machinery and basic materials (on 95% and 99% confidence intervals). The results show that 1% increase in GRP leads to 0.18-0.24% increase in the firm's revenue. HMP arising from the other regions is positive and significant on 99% confidence interval for basic materials and traditional industries, showing that 1% increase in this indicator leads to 0.212-0.438% increase in the firm's revenue. In Table 3.6, HMP measured by urbanization index and by HMP of the other cities is significant for all industries except for integrating machinery (HMP of the other cities is not significant for it).

The effect of regional wage is positive and significant for all industries except for basic materials, where it is not significant. It is the highest for integrating machinery industries, followed by science based industries, traditional industries, and basic machinery industries. In Table 3.6, city wage effect for basic materials was negative. The effect of automobile road density is not significant for all industries considered separately in this specification. The effect of railroad density is negative and significant for traditional and basic materials industries. In Table 3.6 it was positive and significant for basic machinery industries, but in this specification it is not significant for them, though positive. Regional business environment risk is found to produce negative effect on integrating machinery, basic materials and science based industries (in Table 3.6 this coefficient for integrating machinery was not significant). For the other industries it is insignificant, though with negative sign.

In Table 3.8 below, the results for all tradables in specifications with HMP and wage measured on the city level and on the regional level, are summarized.

Table 3.8. Regression results with HMP based on city and regional data: comparison
Dependent variable: ln(*revenue*)

	(1) All tradables-1	(2) All tradables-2	(3) All tradables-3
lfixed_assets	0.234***	0.236***	0.200***
	(22.59)	(22.68)	(28.02)
llabour	0.428***	0.431***	0.516***
	(32.00)	(32.07)	(43.78)
core	9.574***	8.738***	6.757***
	(15.31)	(14.30)	(27.06)
core2	-6.628***	-6.258***	-4.503***
	(-8.20)	(-7.79)	(-17.33)

lloc	0.0312***	0.0287***	0.0202***
	(4.42)	(4.01)	(3.70)
lloc2	-0.00126***	-0.00106***	-0.000652**
	(-3.37)	(-2.77)	(-2.33)
lurbpq	0.282***		
	(10.69)		
ldivpq		0.565***	0.482***
		(5.15)	(13.67)
lgrp			0.200***
			(5.76)
lhmecity	0.451***	0.585***	
	(8.44)	(11.53)	
lhmeR			0.154***
			(4.20)
lwagecity	-0.000303	0.118***	
	(-0.01)	(2.86)	
lwageR			0.306***
			(10.83)
la_road	0.0445*	0.0652**	-0.0431*
	(1.66)	(2.39)	(-1.76)
lrw_road	-0.466	-0.500	-0.171**
	(-1.54)	(-1.57)	(-2.40)
lbusnenvrisk	-0.00805	-0.0166***	-0.0137***
	(-1.53)	(-3.12)	(-2.94)
_cons	-3.588*	-0.0152	6.632***
	(-1.90)	(-0.01)	(16.30)
firm fixed effects	yes	yes	yes
N	49779	49779	96869
adj. R^2	0.428	0.425	0.561
sigma_u	1.153	1.172	1.069
sigma_e	0.522	0.523	0.575
rho	0.830	0.834	0.776

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.8 shows that the coefficients of variables reflecting inputs (labour and capital) and agglomeration variables ('core', localization, diversity) stay robust to changes in specification. Coefficients of agglomeration variables slightly decrease when HMP is measured on the regional level. The impact of HMP measured on the city level (variables 'lurbpq' and 'lhmecity') is relatively higher than the impact with HMP measured on the

regional level (lgrp and lhmeR). The coefficient of wage in specification (1) is not significant. In Models (2) – (3) it is positive and significant, being higher when HMP and wage are measured on the regional level. The difference between Models (1) and (2) is that in Model (1) urbanization coefficient is included into specification, whereas in Model (2) diversity coefficient is included, while urbanization coefficient is not.

The coefficients of variables reflecting the road density are negative and significant with HMP measured at a regional level. However, the coefficient of the variable reflecting automobile roads density is significant and positive when HMP is measured on the city level. It can possibly mean that transport infrastructure is very important for enterprises, but that roads, especially railroads, are located in an inefficient manner. It is probably due partly to the large territory and uneven population density of the country; presence of the railroads and the need for them do not coincide in some territories, so the economic activity can be high in places with not enough railroads. The map of the railroads and cities is presented in Chapter 2, Figure 2.8. The impact of regional business environment risk is negative and significant in specifications (2) and (3), i.e. when HMP is measured both on city and regional levels.

To conclude, the results do not change much with the change of HMP and wage measurement. However, keeping in mind large territories of the regions, measurement on city level is probably more precise, though geographical distances do not reflect the actual roads between cities, which is certainly an important drawback. Below, the results are discussed both for HMP measured on city level and on regional level.

Differences in agglomeration economies between 'young' and 'old' firms. Based on the product cycle theory, it can be assumed that for the 'young' firms[43], urbanization and diversity economies are more important, while for the 'old' firms, localization economies and 'core' variable are more important. In Table 3.9 and Table 3.10 below there are estimates of agglomeration economies for the 'young' and 'old' enterprises separately.

Table 3.9. Regression results for the 'old' and 'young' firms; HMP based on city level data

Dependent variable: ln(*revenue*)

	(1) All tradables-1	(2) All tradables-2	(3) 'Old' firms-1	(4) 'Old' firms-2	(5) 'Young' firms-1	(6) 'Young' firms-2
lfixed_assets	0.234***	0.236***	0.233***	0.234***	0.236***	0.237***
	(22.59)	(22.68)	(15.81)	(15.86)	(16.16)	(16.23)

[43] The year 1995 is chosen as a threshold for 'young' and 'old' firms: if an enterprise was founded before 1995 it is classified as an 'old' enterprise, if it was founded starting from 1995 it is classified as a 'young' enterprise. This division is based on the fact that 1995 was the last year of check privatization in Russia.

	(1) All tradables-1	(2) All tradables-2	(3) 'Old' firms-1	(4) 'Old' firms-2	(5) 'Young' firms-1	(6) 'Young' firms-2
llabour	0.428***	0.431***	0.437***	0.439***	0.417***	0.421***
	(32.00)	(32.07)	(24.32)	(24.29)	(20.82)	(20.98)
core	9.574***	8.738***	9.132***	8.305***	10.24***	9.428***
	(15.31)	(14.30)	(12.43)	(11.65)	(9.47)	(8.81)
core2	-6.628***	-6.258***	-6.612***	-6.152***	-6.787***	-6.577***
	(-8.20)	(-7.79)	(-5.94)	(-5.59)	(-5.52)	(-5.35)
lloc	0.0312***	0.0287***	0.0254***	0.0223***	0.0399***	0.0389***
	(4.42)	(4.01)	(3.12)	(2.71)	(3.07)	(2.95)
lloc2	-0.00126***	-0.00106***	-0.000874**	-0.000651	-0.00191***	-0.00175**
	(-3.37)	(-2.77)	(-2.07)	(-1.52)	(-2.69)	(-2.41)
lurbpq	0.282***		0.244***		0.338***	
	(10.69)		(7.61)		(7.51)	
ldivpq		0.565***		0.424***		0.824***
		(5.15)		(3.39)		(3.90)
lhmecity	0.451***	0.585***	0.440***	0.534***	0.461***	0.667***
	(8.44)	(11.53)	(6.23)	(7.89)	(5.68)	(8.92)
lwagecity	-0.000303	0.118***	0.0346	0.151***	-0.0468	0.0662
	(-0.01)	(2.86)	(0.65)	(2.76)	(-0.77)	(1.07)
la_road	0.0445*	0.0652**	0.0463	0.0580*	0.0384	0.0780*
	(1.66)	(2.39)	(1.39)	(1.72)	(0.84)	(1.66)
lrw_road	-0.466	-0.500	-0.260	-0.274	-0.937*	-1.030*
	(-1.54)	(-1.57)	(-0.69)	(-0.69)	(-1.81)	(-1.94)
lbusnenvrisk	-0.00805	-0.0166***	-0.0126**	-0.0193***	0.000830	-0.0115
	(-1.53)	(-3.12)	(-2.10)	(-3.19)	(0.08)	(-1.09)
_cons	-3.588*	-0.0152	-3.882	-0.511	-2.163	1.673
	(-1.90)	(-0.01)	(-1.61)	(-0.20)	(-0.70)	(0.53)
firm fixed effects	yes	yes	yes	yes	yes	yes
N	49779	49779	32753	32753	17026	17026
adj. R^2	0.428	0.425	0.424	0.421	0.433	0.430
sigma_u	1.153	1.172	1.080	1.087	1.382	1.423
sigma_e	0.522	0.523	0.484	0.485	0.589	0.590
rho	0.830	0.834	0.833	0.834	0.846	0.853

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Agglomeration effects and HMP proved to be present both for 'young' firms and for the firms established before privatization ('old' firms). However, these effects are stronger for the 'young' firms. Contrary to the assumption made above, 'young' firms benefit relatively more not only from diversity and urbanization economies, but also from localization economies. Share of an industry *j* in the total revenue in a city *z* ('core') has an inverted U shape in all specifications. Localization economies also keep an inverted U shape, except in case for 'old' firms, when urbanization variable is not included, but diversity index is present. Then localization is positive. The fact that localization effects stay positive might be explained by product cycle theory, saying that for 'old' firms presence of the other firms in the same industry is beneficial.

Wage was found to have a positive impact on the 'old' firms' revenue. However, this result is not robust to including urbanization or diversity variable into specification interchangeably, and stays positive only when diversity variable is included. The impact of automobile road density remains positive or insignificant, while the impact of railroad density remains negative or insignificant. The impact of regional business environment risk was found to be sensitive to specification; it is negative for all tradables when diversity, but not urbanization variable is included; it is negative only for the 'old' firms, while for the young firms it is not significant. In Table 3.10 below these results are compared to the specification with regional level HMP and wage.

Table 3.10. Regression results for the 'old' and 'young' firms; HMP based on regional data

Dependent variable: ln(*revenue*)

	(1) All tradables-1	(1) All tradables-2	(2) 'Old' firms-1	(2) 'Old' firms-2	(3) 'Young' firms-1	(3) 'Young' firms-2
lfixed_assets	0.200***	0.200***	0.205***	0.205***	0.195***	0.194***
	(28.02)	(28.00)	(20.64)	(20.64)	(18.74)	(18.66)
llabour	0.516***	0.517***	0.518***	0.519***	0.515***	0.516***
	(43.78)	(43.86)	(33.26)	(33.27)	(28.48)	(28.59)
core	6.757***	6.774***	6.213***	6.210***	7.728***	7.782***
	(27.06)	(26.81)	(21.55)	(21.22)	(16.69)	(16.70)
core2	-4.503***	-4.686***	-4.096***	-4.277***	-5.249***	-5.437***
	(-17.33)	(-17.94)	(-12.87)	(-13.34)	(-11.62)	(-12.02)
lloc	0.0202***	0.0182***	0.00863	0.00656	0.0431***	0.0414***
	(3.70)	(3.33)	(1.41)	(1.07)	(3.97)	(3.81)
lloc2	-0.000652**	-0.000531*	0.000117	0.000253	-0.00217***	-0.00209***
	(-2.33)	(-1.89)	(0.36)	(0.78)	(-4.05)	(-3.90)

	(1) All tradables-1	(1) All tradables-2	(2) 'Old' firms-1	(2) 'Old' firms-2	(3) 'Young' firms-1	(3) 'Young' firms-2
ldivpq	0.482***		0.502***		0.441***	
	(13.67)		(12.01)		(6.85)	
divpq		3.217***		2.992***		3.795***
		(11.91)		(9.53)		(7.29)
lgrp	0.200***	0.247***	0.206***	0.244***	0.171***	0.238***
	(5.76)	(6.96)	(5.07)	(5.89)	(2.60)	(3.52)
lhmeR	0.154***	0.142***	0.133***	0.117***	0.235***	0.237***
	(4.20)	(3.88)	(3.16)	(2.77)	(3.26)	(3.28)
lwage	0.306***	0.275***	0.313***	0.293***	0.267***	0.207***
	(10.83)	(9.33)	(10.07)	(8.97)	(4.42)	(3.33)
la_road	-0.0431*	-0.0383	-0.0349	-0.0358	-0.0579	-0.0396
	(-1.76)	(-1.56)	(-1.17)	(-1.19)	(-1.36)	(-0.92)
lrw_road	-0.171**	-0.150**	-0.0502	-0.0211	-0.299**	-0.287**
	(-2.40)	(-2.11)	(-0.69)	(-0.29)	(-2.48)	(-2.39)
lbusnenvrisk	-0.0137***	-0.0150***	-0.0112**	-0.0125**	-0.0236**	-0.0248**
	(-2.94)	(-3.23)	(-2.14)	(-2.40)	(-2.39)	(-2.51)
_cons	6.632***	4.445***	5.896***	3.761***	7.481***	5.098***
	(16.30)	(10.10)	(13.26)	(7.73)	(10.94)	(6.84)
firm fixed effects	yes	yes	yes	yes	yes	yes
N	96869	96869	65450	65450	31419	31419
adj. R^2	0.561	0.561	0.571	0.571	0.548	0.549
sigma_u	1.069	1.314	1.063	1.323	1.141	1.348
sigma_e	0.575	0.575	0.531	0.531	0.658	0.658
rho	0.776	0.839	0.801	0.861	0.751	0.808

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

With HMP and wage measured on the regional level (Table 3.10 above), 'core' keeps the inverted U shape; its values are slightly lower than in Table 3.9. Localization effects become insignificant for the old firms, which is counterintuitive, and become slightly higher for the young firms. Diversity economies are higher for old firms than they were in Table 3.9, and are lower for young firms, resulting into lower diversity economies for the young firms when diversity coefficient is taken with logarithm. However, in specification without logarithm, diversity economies are higher for the young firms, which is in line with assumption that younger firms benefit more than the old ones from diversity in a city.

Concerning HMP, the coefficient of logarithm of GRP is higher for the old firms, and the coefficient of HMP arising from the other regions, is higher for young firms. These

coefficients are significant on 99% confidence interval. This effect is lower than in specification in Table 3.9. Regional wage has a positive impact on the firm's revenue, the effect for the old firms being larger. Compared to the results in Table 3.9, this effect is larger, and is significant in all specifications. Coefficient of automobile road density became insignificant in Table 3.10 while railroad infrastructure remained negative or insignificant. Unlike the results in Table 3.9, where regional business risk was either negative or insignificant, it is negative and significant in Table 3.10 for all specifications.

Specific features of agglomeration economies for private firms and for foreign firms. It can be assumed that private firms benefit more from agglomeration than all firms on average, because they are probably more efficient. Foreign firms are assumed to benefit relatively more from agglomeration than all private firms on average.

Table 3.11. Regression results for private and foreign firms; HMP based on city level data

Dependent variable: ln(*revenue*)

	(1) All tradables	(2) Private firms	(3) Foreign firms
lfixed_assets	0.234*** (22.59)	0.248*** (19.54)	0.294*** (8.97)
llabour	0.428*** (32.00)	0.410*** (27.06)	0.278*** (8.55)
core	9.574*** (15.31)	10.22*** (14.51)	12.03*** (7.17)
core2	-6.628*** (-8.20)	-7.117*** (-8.10)	-8.960*** (-5.37)
lloc	0.0312*** (4.42)	0.0445*** (5.81)	0.0499** (2.35)
lloc2	-0.00126*** (-3.37)	-0.00194*** (-4.77)	-0.00240** (-2.35)
lurbpq	0.282*** (10.69)	0.295*** (9.78)	0.571*** (3.78)
lhmecity	0.451*** (8.44)	0.503*** (7.92)	0.139 (0.47)
lwagecity	-0.000303 (-0.01)	-0.0609 (-1.28)	-0.0432 (-0.26)
la_road	0.0445* (1.66)	0.0247 (0.82)	0.194** (2.03)
lrw_road	-0.466 (-1.54)	-0.246 (-0.80)	-2.765* (-1.75)

lbusnenvrisk	-0.00805	-0.0120*	-0.0255
	(-1.53)	(-1.88)	(-1.38)
_cons	-3.588*	-5.580***	9.307
	(-1.90)	(-2.81)	(0.93)
firm fixed effects	yes	yes	yes
N	49779	37164	4209
adj. R^2	0.428	0.433	0.384
sigma_u	1.153	1.077	2.116
sigma_e	0.522	0.525	0.610
rho	0.830	0.808	0.923

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The results in Table 3.11 above show that agglomeration effects are stronger for private firms (national and foreign) than for all firms producing tradable goods on average and are stronger for the foreign firms than for all private firms on average. This pattern holds for 'core', localization and urbanization effects. Foreign firms benefit from the highest agglomeration effects possibly because they have a wider choice of location and analyze the possibilities of location better than the national firms do. Besides, it seems that they can benefit more from agglomeration economies due to better management (Vakhitov, 2008).

HMP proved to be stronger for private firms than on average for 'tradables'. For the foreign firms this effect is partly significant, i.e. urbanization level in the city where a firm is located positively affects revenue, and the coefficient is larger than for the other firms. However, HMP arising from the other cities turned out to be insignificant for the firms with foreign ownership. Automobile road density is found to be positive and significant for the foreign firms, while railroad density is found to be negative and significant for them. For private firms both road infrastructure variables are not significant. The risks of business environment are negative and significant for the private firms. In Table 3.12 below these results are compared to the specification with HMP and wage measured on the regional level.

Table 3.12. Regression results for private and foreign firms; HMP based on regional data
Dependent variable: ln(*revenue*)

	(1) All tradables	(2) Private firms	(3) Foreign firms
lfixed_assets	0.200***	0.217***	0.293***
	(28.02)	(24.96)	(11.89)
llabour	0.516***	0.504***	0.364***
	(43.78)	(37.33)	(12.53)
core	6.757***	6.853***	7.747***

	(1) All tradables	(2) Private firms	(3) Foreign firms
	(27.06)	(25.45)	(8.94)
core2	-4.503*** (-17.33)	-4.392*** (-15.86)	-4.772*** (-5.27)
lloc	0.0202*** (3.70)	0.0387*** (6.41)	0.0446*** (2.87)
lloc2	-0.000652** (-2.33)	-0.00168*** (-5.43)	-0.00182** (-2.19)
ldivpq	0.482*** (13.67)	0.534*** (12.60)	0.811*** (4.49)
lgrp	0.200*** (5.76)	0.181*** (4.33)	0.0993 (0.67)
lhmeR	0.154*** (4.20)	0.189*** (4.31)	0.274* (1.75)
lwage	0.306*** (10.83)	0.280*** (8.16)	0.213* (1.73)
la_road	-0.0431* (-1.76)	-0.0484* (-1.74)	0.0580 (0.55)
lrw_road	-0.171** (-2.40)	-0.203** (-2.26)	-0.0389 (-0.18)
lbusnenvrisk	-0.0137*** (-2.94)	-0.0183*** (-3.27)	-0.0254 (-1.54)
_cons	6.632*** (16.30)	7.067*** (13.98)	6.394*** (4.68)
firm fixed effects	yes	yes	yes
N	96869	71169	7156
adj. R^2	0.561	0.557	0.532
sigma_u	1.069	0.994	1.265
sigma_e	0.575	0.584	0.696
rho	0.776	0.744	0.768

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Agglomeration effects are robust to changes made in Table 3.12 above, compared to Table 3.11, i.e. when HMP and wage are measured on the regional level. The results show that foreign firms benefit the most from agglomeration, followed by private firms, which benefit more than all firms producing tradable goods on average. The coefficients of 'core' and localization are slightly lower than in specification in Table 3.11, but keep the inverted U shape and remain significant on 99% confidence interval, and in case of localization for

foreign firms, on 95% confidence interval. Diversity coefficient included in specification of Table 3.12, is substantially larger for foreign firms, followed by private firms, for which it is also higher than on average for all firms producing tradable goods.

On the regional level, HMP (reflected by regional GRP and GRP of the other regions taking into account distances between regions' capitals) proved to be positive and significant for all tradables and for the national firms. It also proved to be partly significant for the firms with foreign ownership, i.e. GRP of the surrounding regions matter, though the coefficient is significant only on 90% confidence interval. GRP of the region where the firm works turned out to be insignificant for the foreign firms. The same time, urbanization (revenue of all enterprises in a city except the firm itself) in Table 3.11 was significant. It confirms that considering city level is probably a correct strategy. Regional wage turned out to be positive and significant for all types of firms. Coefficient for automobile road infrastructure changed the sign and became negative. Coefficient for railroad infrastructure remained negative or insignificant. Regional business risk turned out to be negative and significant again for private firms (both national and foreign firms), but not for the foreign firms considered separately.

Table 3.13. **Regression results for different types of cities; HMP is based on city level data**
Dependent variable: ln(*revenue*)

	(1) All tradables	(2) Agglomeration centers	(3) Other cities within agglomerations	(4) Monotowns	(5) All other cities
lfixed_assets	0.235***	0.242***	0.240***	0.280***	0.216***
	(22.61)	(15.31)	(5.08)	(8.57)	(14.88)
llabour	0.428***	0.381***	0.350***	0.290***	0.513***
	(32.03)	(22.98)	(7.48)	(4.81)	(22.64)
core	9.600***	9.549***	9.670***	4.807**	12.76***
	(15.35)	(8.44)	(7.73)	(2.41)	(14.01)
core2	-6.663***	-13.85***	-8.549***	-0.711	-10.21***
	(-8.22)	(-5.27)	(-4.72)	(-0.41)	(-8.01)
lloc	0.0322***	0.0376***	0.109***	0.0471**	0.0335***
	(4.57)	(2.83)	(4.09)	(2.55)	(3.51)
lloc2	-0.00133***	-0.00131**	-0.00708***	-0.00158	-0.00177***
	(-3.54)	(-2.16)	(-4.82)	(-1.33)	(-3.27)
lurbpq	0.280***	0.265***	0.249***	0.308***	0.286***
	(10.64)	(4.24)	(3.04)	(3.84)	(8.31)
lhmecity	0.493***	0.375***	0.531***	0.439**	0.755***

	(1) All tradables	(2) Agglomeration centers	(3) Other cities within agglomerations	(4) Monotowns	(5) All other cities
	(9.30)	(3.26)	(4.50)	(2.41)	(8.01)
lwagecity	-0.0405	0.0901	0.0144	-0.109	-0.277***
	(-1.00)	(0.97)	(0.43)	(-0.60)	(-3.49)
a_road	0.000347***	0.000282**	0.000237	0.000219	-0.000171
	(3.44)	(2.21)	(0.64)	(0.42)	(-0.67)
rw_road	-0.00214	-0.000929	-0.0154	0.00380	-0.00478***
	(-1.58)	(-0.27)	(-0.74)	(0.83)	(-3.02)
lbusnenvrisk	-0.0109**	-0.0208***	0.0180	-0.0343*	-0.0144
	(-2.10)	(-3.16)	(0.91)	(-1.70)	(-1.29)
_cons	-5.858***	-4.322*	0.541	-5.663*	-9.557***
	(-6.34)	(-1.94)	(0.05)	(-1.96)	(-6.01)
firm fixed effects	yes	yes	yes	yes	yes
N	49866	24458	1872	2578	21393
adj. R^2	0.428	0.401	0.549	0.500	0.462
sigma_u	1.151	1.202	2.896	1.247	1.125
sigma_e	0.522	0.530	0.449	0.446	0.515
rho	0.830	0.837	0.977	0.886	0.827

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Share of an industry *j* in the total revenue in a city *z* ('core' coefficient) is significant for enterprises located in all types of cities. For all types of cities, except for monotowns 'core' has an inverted U shape and is significant on 99% confidence interval. For monotowns it is positive and significant on 95% confidence interval. 'Core' is the highest for 'all other cities', followed by 'other cities within agglomeration', agglomeration centers and monotowns. Localization economies also have an inverted U shape for all cities types, except for monotowns, where they are positive. The significance level is similar to that of 'core'. Localization economies having an inverted U shape are the highest for the firms located in the cities within agglomerations but not agglomeration centers, followed by agglomeration centers and other cities. Positive effect of 'core' and localization show that firms' concentration in the same industry is particularly important for the enterprises in monotowns.

Urbanization economies are positive and significant for all types of towns, on 99% confidence interval. They are the highest also for monotowns, followed by 'all other cities', agglomeration centers and other cities within agglomeration. The impact of HMP arising from other cities is the highest for 'all other cities', followed by 'other cities within agglomerations', then by monotowns and agglomeration centers.

Towns, including the monotowns, located within agglomerations, are assumed to have better access to transport infrastructure and shorter distances to the other towns, implying more opportunities to benefit from larger HMP of the surrounding cities than towns outside of agglomeration. However, the results emphasize the importance of HMP for the towns located outside agglomerations. As revenue generated in these towns is lower than in the agglomeration centers, for them HMP coming from the other cities is relatively more important than it is for the agglomeration centers. Keeping in mind that the variable *lhmecity* that reflects HMP based on city data is constructed using physical distance, it really is market *potential*, that can be realized only if road infrastructure allows it. This fact draws our attention to the importance of road infrastructure development.

Wage effect is either insignificant or negative (for 'all other cities'). The effect of automobile road density is positive in case of agglomeration centers and is not significant for other cities. Railroad density coefficient is either insignificant or negative (for all other cities). Regional business environment risk is significant and negative for agglomeration centers and monotowns, and is not significant for other cities.

In the analysis carried out in this chapter, it was assumed that endogeneity caused by simultaneity is not present. In order to check this assumption, we used lagged agglomeration variables as instruments for the agglomeration variables, and applied 2SLS and GMM interchangeably. The results remained robust (they are not presented here). In the next section we address the question if localization level is optimal in Russia.

3.5. Is localization level optimal in Russia?

Below distribution of enterprises according to the localization levels in the cities and industries where they operate, is presented. The econometric estimation results suggest an inverted U shape of localization economies, implying the existence of optimal localization level. To understand better the motivation of cluster policies, it is important to find out if the actual firms' distribution among cities maximizes localization economies. The main argument behind cluster policies is that actual localization levels are sub-optimal as localization economies are not always internalized by firms, as well as the other externalities in the market economy (Martin et al., 2011). To check the correspondence of the actual distribution of firms to the optimal one, localization economies surplus was calculated based on the formula:

$$LE_{jz} = \hat{\beta}_1 loc_{jz} + \hat{\beta}_2 loc_{jz}^2 . \qquad (3.2)$$

In equations (3.3) – (3.4) $\hat{\beta}_1$ and $\hat{\beta}_2$ are regression (3.1) estimates for all enterprises working in the industries that produce tradable goods. Equation (3.3) is for Model 1, Table 3.4, specification with *diversity* index:

$$LE_{jz} = 0.0287 loc_{jz} - 0.00106 loc_{jz}^2, \quad (3.3)$$

In equation (3.4) $\hat{\beta}_1$ and $\hat{\beta}_2$ are regression (3.1) estimates for Model 2, Table 3.4, specification with *urbanization* index (this specification is used further in this chapter):

$$LE_{jz} = 0.0312 loc_{jz} - 0.00126 loc_{jz}^2, \quad (3.4)$$

Theoretical localization economies, received from the equation (3.4), are plotted against possible localization levels. The actual distribution of firms along the localization levels is received applying kernel density function. The two lines are depicted in Figure 3.1 below.

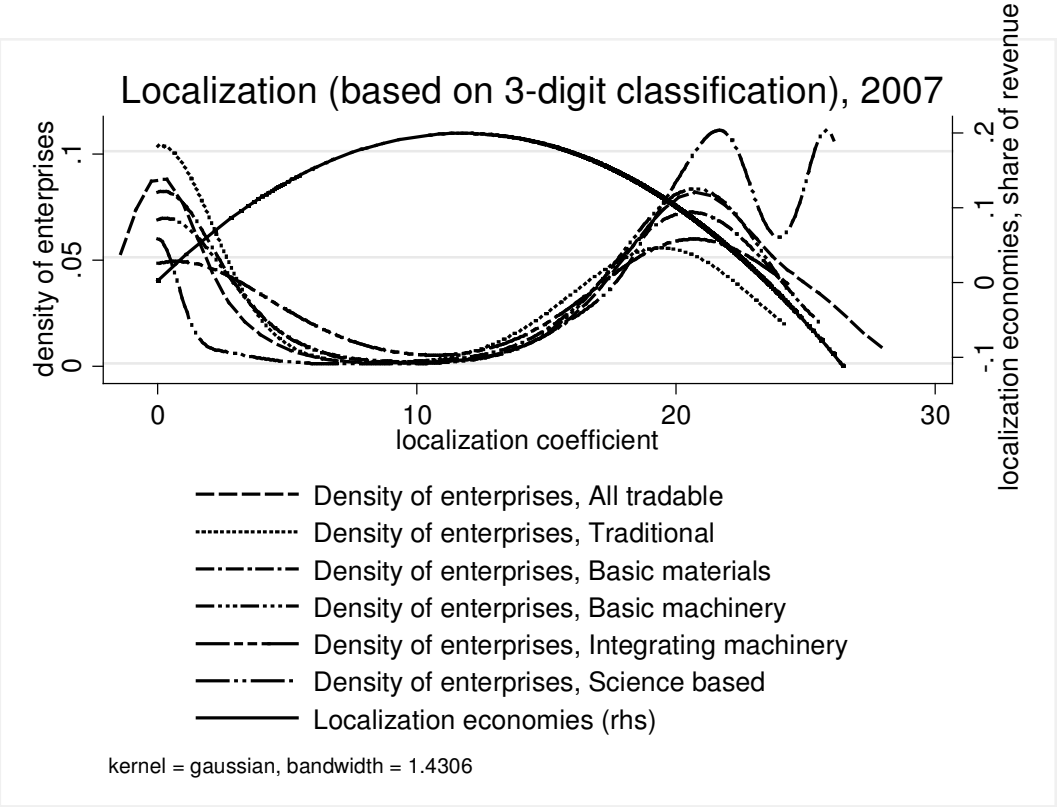


Figure 3.1. Actual distribution of firms along the localization levels vs. theoretical curve of localization economies

For 3-digit definition of industry localization, the effect is positive for the range of localization from 0 to 56.8 bn roubles (according to equation 3.4). At the same time, there is maximum of localization economies at 12.38 - 13.53 or 238,220 – 757,467 roubles, equation

(3.4) giving lower number. The results do not provide exact understanding of the localization optimum. However, it was found that starting from even low level of revenue generated by the other firms in the industry, localization economies are positive, and they remain positive for the most enterprises in the sample (Figure 3.1 above) until rather high localization level is reached. In equation (3.3.) this level is 0.57 bn roubles, while in specification using HMP for regions, maximum localization level with positive localization effects was 28.5 bn roubles. Vorobyev et al. (2010) concluded that any localization level up to 17.5 bn is better than zero level, but levels higher than 130 mn roubles are not necessary.

As for the results with HMP for regions, localization effect for science based industries is positive; this result is comparable with the result received by Henderson (2003), who concludes that the number of other own industry plants positively affects productivity in high tech industries (but not in machinery industries). For the 'old' firms in specification with HMP on the city level and with diversity index, localization economies are also positive, while in specification with urbanization index they have an inverted U shape with optimum at 2 mn roubles (Models 4 and 3 respectively in Table 3.9).

On the graph, zero localization level means that there is only one firm in a city; such situation can occur in monotown. Compared to the theoretical optimum, in the actual firms' distribution along localization levels, there is considerably large number of firms in the cities with very small localization level, and with the localization level higher than optimal. The same time, not so many enterprises are located in the cities with optimal localization level. Therefore, on one hand, there are firms located in smaller towns where competition is not so intensive. For them, incentives and opportunities for development are low. These firms are a reasonable focus for policy aimed at attracting firms in cities where localization is less than optimal. Such policy can include improvement of business climate and facilitating agglomeration with neighbouring territories, particularly, via improvement of transport infrastructure.

On the other hand, enterprises located close to optimal localization levels internalize localization economies; for them cluster policy would not be helpful (Vorobyev et al., 2010). There are also enterprises located on the localization levels exceeding the optimal one, i.e. on the levels where the localization impact on firm performance is still positive but decreasing. Diminishing localization effects might arise mainly because of competition for resources as we consider tradable industries. However, there might also be competition for consumers. To decrease these effects, it seems helpful to improve opportunities for trade with other cities, regions, and for international trade. It involves improvements of official regulations, business climate, business services, and transport infrastructure.

The sample under consideration includes firms created under centrally-planned decision-making. It was assumed that for them simultaneity problem should not arise because their location was not likely to be motivated by efficiency. Simultaneity can occur when more ex-ante productive firms choose a certain location. As a result, there may be a large number of productive enterprises in this location, which would create an illusion that it is due to the location's characteristics, such as agglomeration level, market potential etc. However, it can be specific enterprise's factors that are actually behind its successful performance. Such correlation between productivity and location decisions is typical for profit-maximizing firms, but the firms under central planning were not maximizing profit. Therefore for the 'old' firms simultaneity problem is not likely to exist. Moreover, from Figure.3.2 below it is evident that distribution of 'old' and 'young' enterprises along localization levels is very similar. Therefore we will suppose that simultaneity problem does not affect the results for the whole sample of firms. The same conclusion was made by Vorobyev et al. (2010) based on comparing distribution of firms using enterprise level database for 2001-2005.

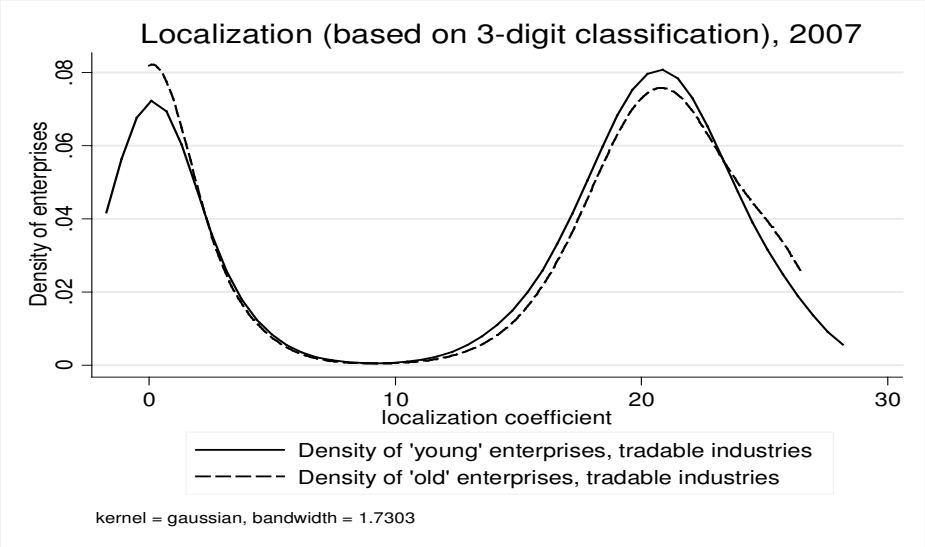


Figure 3.2. Actual distribution of 'old' and 'young' firms by localization level

From Figure 3.2 it can be concluded that 'young' enterprises also do not demonstrate behavior leading to endogeneity. In fact, a lot of 'young' firms are based on production facilities of the enterprises founded during the times of central planning. Within this logic it is assumed that simultaneity problem is not present in estimation done in this chapter and dealing with the impact of agglomeration on enterprise performance. The same time, the distribution of national and foreign enterprises differs (Figure 3.3 below). Therefore agglomeration indices associated with foreign firms are likely to be endogeneous due to simultaneity, and instruments are applied to deal with this problem in the next chapter, following the approach of Martin et al. (2011).

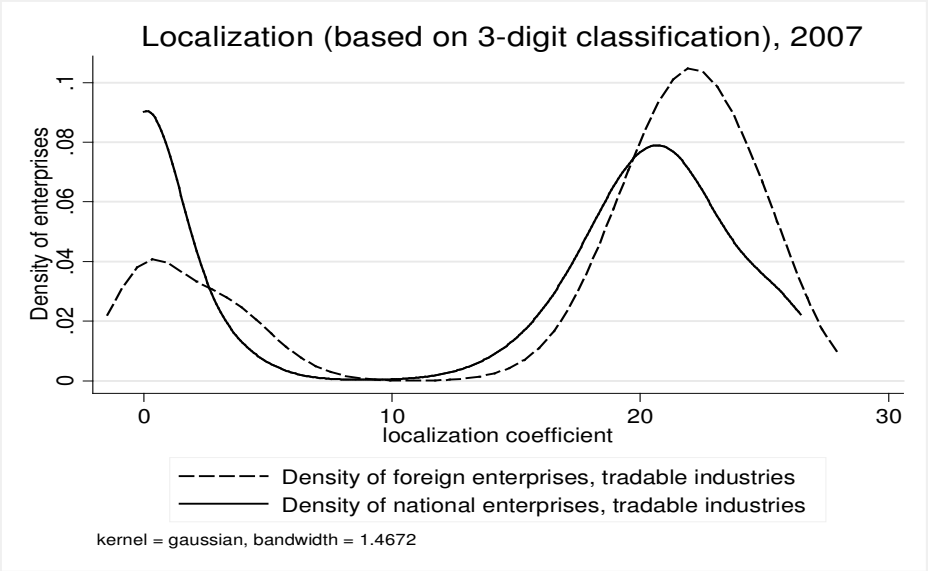


Figure 3.3. Actual distribution of national and foreign firms by localization level

3.6. Conclusion

In Chapter 3, the impact of agglomeration and home market potential on the enterprise performance was analyzed. The first hypothesis that both urbanization and HMP have a positive impact on the enterprise performance was confirmed. The second hypothesis that localization economies have an inverted U shape was confirmed for the majority of specifications. The same is true about 'core'. Urbanization economies are positive on 99% confidence interval and are robust to specification changes. Diversity economies are positive as well. HMP is significant and positive in the majority of specifications. As it is related to distance and transport costs, it might be concluded that transport infrastructure improvement and decrease in communications costs would increase the effects arising from HMP.

The third hypothesis that regional transport infrastructure and business climate are the factors essential for enterprise performance was partly confirmed. Concerning road infrastructure, automobile road density proved to be positive and significant in some specifications, while railroad density turned out to be either insignificant or negative, possibly meaning that businesses with relatively higher revenues are located in places where railroad density is relatively lower. Positive sign of automobile roads density emphasizes the importance of road infrastructure for enterprise performance. Regional business risk measured by the Analytical agency 'Expert' has a negative sign or is insignificant. It draws our attention again to the importance of business climate for successful work of enterprises. Results proved to be largely robust to measuring HMP and wage on city and regional levels.

The fourth hypothesis that human capital level is important for enterprise performance was not confirmed. However, this outcome seems to be caused by the choice of wage as a proxy for human capital quality. On one hand, higher wage means higher quality of human capital leading to higher productivity, on the other hand, costs for enterprises arising from higher city wage might outweigh the first effect. The effect of wage on productivity is controversial: in some specifications it is positive, and in the others it is negative or insignificant.

Answering the question 'Is theory confirmed by data?', empirical analysis confirms the existence of positive agglomeration effects and positive effect of HMP. Data also reflects the congestion attained at some point of increasing agglomeration, and leading to the decrease in agglomeration economies: localization economies have an inverted U shape.

The next question was 'How do these effects vary across cities of different type, across industries and types of enterprises?' Concerning the groups of industries, share of an industry j in the total revenue in a city z ('core' coefficient) turned out to be significant for all industries. It has an inverted U shape and is relatively more important for traditional industries, followed by basic machinery industries, basic materials, integrating machinery and science based. Localization coefficient has an inverted U shape; it is the largest for integrating machinery, followed by basic materials and basic machinery. With HMP and wage measured on regional level, localization coefficient has an inverted U shape for all industries except for science based industries, for which it is positive and significant. Diversity is significant and positive, showing that 1% increase in diversity leads to 0.48% increase in enterprise revenue. Urbanization is significant and positive too, 1% increase in urbanization leading to 0.27% increase in enterprise revenue.

We also posed a question 'Which kinds of enterprises gain more from agglomeration effects - national or foreign ones; private or public; created before or after the privatization, more productive or less productive, etc.'? Agglomeration effects proved to be higher for the foreign firms than for all firms belonging to tradable industries on average. For private firms (national or foreign) these effects are also stronger than on average for all tradables, but lower than for the foreign firms. Localization, urbanization and diversity effects proved to be almost twice higher for the foreign firms than on average for all tradables; the effect arising from the industry share in the city revenue is also higher for them.

As for different types of cities, positive effect of 'core' and localization show that concentration of the firms in the same industry is particularly important for the enterprises in monotowns. Urbanization economies are also found to be the highest for monotowns. It might imply the importance both of further specialization and of development of business in

industries other than industries of specialization there. The degree of further specialization and development of the other sectors certainly depends on situations in particular towns. The sample under consideration includes only manufacturing industries; services are not included. However, development of services, including business services seems important in all cities to sustain development of manufacturing. Development of small business could be a way to increase 'urbanization' and 'diversity' economies. The results emphasize the importance of HMP arising from the other cities, primarily for the towns located outside agglomerations, followed by the towns within agglomerations which are not agglomeration centers, monotowns and agglomeration centers.

Finally, the following question was posed: 'Do enterprises take into account agglomeration economies?' As for optimal localization level, there are firms located in smaller towns where competition is not so intensive and opportunities for development are low. These firms are a reasonable focus for policy, such as improvement of business climate and facilitating agglomeration with neighbouring territories, particularly, via improvement of transport infrastructure. On the other hand, enterprises located close to optimal localization levels internalize localization economies; for them cluster policy would not be helpful. There is also a group of enterprises that face diminishing localization effects. For them, openness to trade with the other cities, regions, and for international trade seems to be relevant. When the model was estimated with instruments using 2SLS and GMM methods, the results remained robust, though there were slight changes (the results are not included in the chapter).

4. The impact of foreign direct investment concentration on enterprise performance

4.1. Introduction

In the previous chapter the effects of spatial concentration of economic activity were analyzed. In this chapter these effects are subdivided into those arising from national and foreign firms' concentration in cities. Agglomeration process and foreign direct investment (FDI) inflow mutually affect each other, i.e. the determinants and effects of agglomeration processes are interrelated with the determinants and effects of FDI. 'Circular logic' of FDI implies that higher level of economic growth and development makes a country able both to attract FDI and to benefit from it due to better absorptive capacity. Further, increase in FDI inflow stimulates economic development of a country (Ledyaeva et al., 2010). Then the process continues. It is the same kind of logic as that typical for agglomeration effects.

Overall, both agglomeration effects and FDI stimulate economic growth, i.e. growth of output, revenue, wage, number of workers, in the cumulative process.

Bode et al. (2009) cite the evidence from a number of papers that the majority of foreign firms in the USA choose relatively advanced states, and locate where they can benefit from agglomeration externalities. In turn, FDI concentration within the USA may have affected spatial density of business activity explaining a significant part of the productivity variation across US states. In terms of home market potential and FDI, critical level of demand attracts FDI, then FDI creates new demand, furthermore, it creates demand for specific goods or services, such as high quality business services.

In fact, keeping in mind microfoundations of agglomeration effects considered in literature review (input sharing, matching and knowledge spillovers), it becomes evident that agglomeration and FDI effects are interrelated. Firstly, *input sharing* implies benefits from infrastructure and inputs attracting both national and foreign firms. Secondly, *matching mechanisms* include easier search, particularly on the labour market, that again attract both national and foreign firms. Thirdly, *knowledge spillovers* imply exchange of information arising from proximity to the other firms. These effects arising from business agglomeration are interrelated with FDI spillovers, as they include effects generated by foreign firms. For example, FDI is associated with enhanced quality of intermediate goods. Higher quality of intermediate goods attracts national firms and firms with FDI producing final goods. It again generates positive agglomeration effects, further attracting FDI. Overall, it can be assumed that FDI serves as catalyst of increased economic activity in a city (agglomeration); on the other hand, agglomeration further attracts economic activity, including firms with FDI. The same logic can be employed to consider possible negative effects or congestion costs of FDI and agglomeration.

The objective is to reveal the impact of business activity concentration on enterprise productivity, paying special attention to the effects arising from concentration of national and foreign firms. Firms with FDI or foreign firms here are the firms with foreign ownership. The firm level and city level data is employed, taking into account such regional characteristics as road infrastructure and business environment risks. We look at the effects of enterprise concentration for different groups of industries and cities, examine HMP, and look at the role of infrastructure and institutions. Questions posed in this chapter are the following. What are the effects on enterprise performance generated by concentration of national and foreign enterprises? How do these effects depend on the enterprise characteristics, industry, and city type?

The first hypothesis is that urbanization effects arising from national firms' concentration and HMP have a positive impact on the enterprise productivity. The second hypothesis is that localization effects have an inverted U shape, due to congestion costs and crowding out effect of competition that might prevail at a certain point of concentration. The third hypothesis is that FDI concentration may have controversial localization effects due to competition in the same industry, but positive urbanization effects. The other two hypotheses are the same as in the previous chapter. The fourth hypothesis states, that regional transport infrastructure and business climate are important for enterprise performance. The fifth hypothesis is that human capital level is also important for enterprise performance.

To test these hypotheses, we analyze how firms' revenue is affected by agglomeration of national and foreign firms, HMP, city average wages, transport infrastructure and business environment, controlling for the firms' inputs (labour and capital). The econometric model is based on the approach by Martin et al. (2011), Henderson (2003), Vakhitov (2008), Vorobyev et al. (2010). Ideas from the papers mentioned below are also taken into consideration.

There is vast literature studying FDI effects at the country and regional levels (Navaretti and Venables (2004), Castelani and Pieri (2010), etc.). Theoretical foundations and results of empirical research on agglomeration and FDI effects are summarized in the literature review (Chapter 1). Effects from FDI include the effects on enterprise productivity, infrastructure, particularly, financial infrastructure, on quality of labour force, and on variety of resources. However, the nature and extent of these effects seem to depend on motivation of multinational corporations (MNCs) that are involved in FDI. MNCs' motivation is associated with their strategies that can be traditional (market seeking, asset seeking, efficiency seeking, natural resources seeking) or more complex global strategies (Andreff, 2003). MNCs' strategies are considered in more detail in Chapter 5. Influence of foreign firms on all the firms in a city probably depends on the prevalence of a certain strategy. For instance, it can be assumed that it is less productive firms that are interested in lower wages i.e. pursue efficiency seeking strategies. Then, there is a question, how much these firms can contribute to improvements in production and management.

Besides depending on MNCs' strategies, indirect effects from FDI also depend on the national firms' characteristics (Castelani, Pieri, 2010). Potential spillovers take place only if national firms are able and have motivation to invest in absorbing foreign technologies and skills (Blomstrom and Kokko, 2003). Firms' productivity differs substantially within industries. Productivity of national firms seems to reflect capability to adopt new technologies, i.e. the absorptive capacity. There is a threshold productivity level, which makes firms capable of adopting new knowledge. When foreign firms enter a local market, national

firms with lower productivity are likely to leave the market as they have losses, while firms with higher productivity will stay and benefit from presence of foreign firms; those of them with relatively higher productivity will enter a foreign market themselves.

Literature review for this chapter is provided in Chapter 1. The next section is devoted to the econometric model specification. In section 4.3 data and summary statistics are discussed. Section 4.4 is devoted to estimation issues. Section 4.5 discusses the results, and the conclusions follow.

4.2. Econometric model

The model analyses factors on industrial, city and regional levels affecting enterprise productivity. Industrial level factors are activities of national and foreign firms in an industry in a city. Concerning the effects generated by foreign firms, presence of FDI is associated with infrastructure development, larger number of workforce in an industry in a city and larger demand for final goods. FDI is assumed to enhance productivity by stimulating development within an industry through direct effects, spillovers and competition.

Agglomeration variables are subdivided into those generated by national and foreign firms. Factors considered on the city level are activity of national and foreign firms in the same industry where a firm works (localization), activity of national and foreign firms in various industries other than that where a particular firm works (urbanization), HMP and wage. On the regional level, we consider density of regional road infrastructure and business environment risks. Estimation is carried out by regressing the logarithm of firm's revenue on the firm's characteristics, indices of agglomeration, city wage and regional characteristics; enterprise level fixed effects are used. Regional transport infrastructure and business environment risks are included into the regression; the other regional factors are reflected by the enterprise fixed effects, as in the sample enterprises do not change regions. The following equation is estimated:

$$\begin{aligned}\ln(revenue)_{ti}^{jz} &= \beta_0 + \beta_1 \ln(\text{capital})_{ti}^{jz} + \beta_2 \ln(\text{labour})_{ti}^{jz} + \beta_3 core_t^{jz} + \beta_4 \left[core_t^{jz} \right]^2 + \beta_5 \ln\left(locnl\right)_{ti}^{jz} + \\ &+ \beta_6 \left[\ln\left(locnl\right)_{ti}^{jz} \right]^2 + \beta_7 \left[\ln\left(locf\right)_{ti}^{jz} \right] + \beta_8 \ln\left(\text{urbnl}\right)_t^{jz} + \beta_9 \ln\left(\text{urbf}\right)_t^{jz} + \beta_{10} \ln\left(hmp\right)_t^{z} + \\ &+ \beta_{11} \ln\left(wage\right)_t^{z} + \beta_{12} a\_road_t^r + \beta_{13} rw\_road_t^r + \beta_{14} \ln\left(\text{busnenvrisk}\right)_t^r + \varphi_i + \varepsilon_{ti},\end{aligned} \tag{4.1}$$

where *j* is industry index, *z* is city index, *r* is regional index, and *i* is firm index. In Table 4.1 below there are definitions of variables used in the econometric analysis in this chapter.

Table 4.1. Definition of variables used in the econometric analysis

Variable	Definition
Enterprise level characteristics.	
$\ln(revenue)_{ti}^{jz}$	Logarithm of enterprise revenue (revenue is in roubles)
$\ln(\text{capital})_{ti}^{jz}$	Logarithm of fixed assets (fixed assets are in roubles)
$\ln(\text{labour})_{ti}^{jz}$	Logarithm of labour force (number of employees)
Agglomeration indices[44]. City level.	
$\ln\left(loc\right)_{ti}^{jz}$	Localization coefficient (Equation (2.1))
$\ln\left(locnl\right)_{ti}^{jz}$	Localization coefficient based on national firms revenue
$\ln\left(locf\right)_{ti}^{jz}$	Localization coefficient based on foreign firms revenue
$core_t^{jz}$	A share of an industry *j* in the total revenue in a city *z*
LQ_{jz}	Localization quotient (Equation (2.2))
$\ln\left(divpq\right)_t^z$	Diversity coefficient (Equation (2.7.))
$\ln\left(\text{urb}\right)_t^{jz}$	Coefficient of urbanization (Equation 2.3))
$\ln\left(\text{urbnl}\right)_t^{jz}$	Urbanization coefficient based on national firms revenue
$\ln\left(\text{urbf}\right)_t^{jz}$	Urbanization coefficient based on foreign firms revenue
Home market potential. Regional level: HMPregion= $\ln\left(grp\right)_t^r + \ln\left(hmp\right)_t^r$.	
$\ln\left(grp\right)_t^r$	Logarithm of gross regional product (GRP is in mn roubles)
$\ln\left(hmp\right)_t^r$	The sum of gross regional products of other regions divided into the distances between the region *R* and the other regions. $hmp_t^R = \sum_{r \neq R} \frac{grp_t^r}{dist_{R,r}}$, where $dist_{R,r}$ are the *railroad* distances between the regional capitals
Home market potential. City level: HMPcity= $\ln\left(urbnl\right)_t^{jz} + \ln\left(urbf\right)_t^{jz} + \ln\left(hmpcity\right)_t^z$.	
$\ln\left(hmpcity\right)_t^z$	Total revenue in all industries in all cities other than the city where the firm is located, divided into the distances between the city *Z* and the other cities: $hmpcity_t^Z = \sum_{z \neq Z} \frac{revenue_t^z}{dist_{Z,z}}$, where $dist_{Z,z}$ are the *physical* distances between the cities
Regional transport infrastructure.	
$\left(\text{a\_road}\right)_t^r$	Density of auto roads in region r, end of year, km of roads per 1000 sq km of territory
$\left(rw\_\text{road}\right)_t^r$	Density of railroads in region r, end of year, km of roads per 1000 sq km of territory

[44] All coefficients ('core', localization quotient, localization index, diversity and urbanization indices) are measured based on 3-digit level of OKVED classification

Variable	Definition
Regional investment climate (Indices by analytical agency 'Expert').[45]	
$\ln\left(\text{expertinvptl}\right)_t^r$	Investment potential in region r at time t (Regional investment potential by the Analytical agency Expert)
$\ln(\text{busnenvrisk})_t^r$	Business environment risks, region r, time t (Regional investment risk by the Analytical agency Expert)
Other regional characteristics.	
$\ln\left(\text{wageR}\right)_t^r$	Average monthly nominal wage of organizations' employees, payroll
Other city characteristics.	
$\ln\left(\text{wagecity}\right)_t^z$	Average monthly nominal wage, payroll, roubles

$\ln\left(div\right)_t^z$ is used interchangeably with $\ln\left(\text{urb}\right)_t^{jz}$, as cities with higher urbanization tend to be more diversified. Detailed explanation of indices and other variables is in Chapters 2 and 3.

4.3. Data and descriptive statistics

In the Tables 4.2 and 4.3 below, descriptive statistics of the variables used in the analysis is presented.

Table 4.2. Descriptive statistics for tradable industries; firms, for which *region* level data is present

Variable	Obs	Mean	Std.Dev.	Min	Max
revenue	119383	4.37e+08	3.50e+09	57.00	3.12e+11
fixed_assets	117026	1.34e+08	2.25e+09	16.00	6.02e+11
labour	112925	490.67	1,584.66	1.00	84,918.00
lrevenue	119383	17.37	2.36	4.04	26.47
lfixed_ assets	117026	15.76	2.61	2.77	27.12
llabour	112925	4.99	1.55	0	11.35
core	127712	0.08	0.18	0	1.00
core2	127712	0.04	0.14	0	1.00
llocnl	127712	11.18	10.13	0	26.59
llocnl2	127712	227.57	221.91	0	707.21
llocf	127712	4.77	8.76	0	25.46
divpq	127712	0.32	0.27	0.01	0.84
ldivpq	127712	-1.66	1.16	-4.99	-0.18
ldivpqnl	127559	-1.70	1.17	-4.99	-0.19
ldivpqfdi	80959	-2.81	1.62	-5.01	-0.48

[45] Details are in Appendix C 'Data sources'

Variable	Obs	Mean	Std.Dev.	Min	Max
lurbnl	127712	22.63	4.66	0	27.92
lurbf	127712	6.49	10.94	0	27.39
lhmecity	127712	22.87	0.67	20.39	26.82
lwagecity	60113	9.15	0.57	7.46	12.05
a_road	127622	181.23	131.66	0.80	536.00
rw_road	127622	262.72	181.85	0	583.00
lbusnenvrisk	127712	2.85	1.17	0	4.48
private	127712	0.74	0.44	0	1
state	127712	0.38	0.49	0	1
fdi	127712	0.08	0.27	0	1
new	127712	0.36	0.48	0	1
productivity	112374	950,861.69	2.39e+07	1.17	4.02e+09
highprod	127712	0.23	0.42	0	1
highprodmed	127712	0.56	0.50	0	1

In Table 4.3 below there is descriptive statistics for the sample with city level data.

Table 4.3. Descriptive statistics for tradable industries; firms for which *city* level data is present

Variable	Obs	Mean	Std.Dev.	Min	Max
revenue	56705	5.66e+08	3.54e+09	1,000.00	1.83e+11
fixed_assets	55479	1.40e+08	9.78e+08	23.00	5.49e+10
labour	53360	526.23	1,651.34	1.00	81,000.00
lrevenue	56705	17.86	2.26	6.91	25.94
lfixed_ assets	55479	16.03	2.58	3.14	24.73
llabour	53360	5.11	1.50	0	11.30
core	60113	0.04	0.10	0	1.00
core2	60113	0.01	0.07	0	1.00
llocnl	60113	14.76	9.59	0	26.54
llocnl2	60113	309.82	220.72	0	704.14
llocf	60113	6.87	9.97	0	25.43
divpq	60113	0.43	0.25	0.01	0.84
ldivpq	60113	-1.05	0.71	-4.96	-0.18
ldivpqnl	60113	-1.09	0.72	-4.96	-0.19
ldivpqfdi	50295	-2.60	1.60	-4.97	-0.48
lurbnl	60113	24.44	3.76	0.00	27.92
lurbf	60113	9.19	12.30	0.00	27.39

Variable	Obs	Mean	Std.Dev.	Min	Max
lhmecity	60113	23.02	0.53	21.05	24.57
lwagecity	60113	9.15	0.57	7.46	12.05
a_road	60050	207.57	145.39	5.50	536.00
rw_road	60050	287.72	187.67	0	575.00
lbusnenvrisk	60113	2.77	1.13	0	4.47
private	60113	0.75	0.43	0	1
state	60113	0.37	0.48	0	1
fdi	60113	0.09	0.28	0	1
				0	1
new	60113	0.36	0.48	0	1
productivity	53159	1.35e+06	3.36e+07	1.17	4.02e+09
highprod	60113	0.27	0.44	0	1
highprodmed	60113	0.68	0.47	0	1

In Table 4.4 below, the distribution of firms across various types of cities is presented.

Table 4.4. Distribution of firms across city types, tradable industries, 2007

City type	All tradables	Foreign firms	Private firms	State firms
Agglomeration center	4351	546	3128	1742
Within agglomeration; not center	702	69	533	260
Monotown within agglomeration	56	10	48	12
Monotown outside agglomeration	693	58	561	208
Other cities	7039	354	5208	2657
All cities	**12841**	**1037**	**9478**	**4879**

Regional level data is available for 12831 out of 12841 observations. In Table 4.5 below, there are observations, for which city level data is available.

Table 4.5. Distribution of firms across city types, tradable industries, observations with city data, 2007

City type	All tradables	Foreign firms	Private firms	State firms
Agglomeration center	4080	536	2930	1631
Within agglomeration; not center	325	25	247	124
Monotown within agglomeration	20	2	17	4
Monotown outside agglomeration	384	36	326	96
Other cities	3760	164	2884	1304
All cities	**8569**	**763**	**6404**	**3159**

The database includes 8 - 9% of foreign firms. There are 19421 firms in the sample; 18093 of them are national firms and 1328 are firms with foreign ownership. Foreign firms are present in 305 cities out of 1027 cities, for which enterprise level data is available (the number does not change over the years 1999-2008). Concerning cities with population over 100 thousand people, for which city level data is available, the sample covers 166 cities in 2002, and 154 cities in 2008, out of 172 cities, with a tendency of slight decrease over the years 2002-2008 (which might be purely a feature of the available data). Foreign firms are present in 114 cities in 2002, and 106 cities in 2008, also with a tendency of decrease. In Table 4.6 below there is average level of diversity, localization and urbanization for various types of firms (average for the years 1999-2008).

Table 4.6. Average level of agglomeration for various types of firms; all firms

	Mean (divpq)	**Mean (llocf)**	**Mean (llocnl)**	**Mean (lurbf)**	**Mean (lurbnl)**
'old' firms	0.3	4.41	11.56	5.8	22.6
'young' firms	0.28	3.99	11.14	5.68	22.23
firms without private ownership	0.26	3.83	10.51	4.85	22.22
firms with private ownership	0.3	4.44	11.8	6.16	22.57
firms without state ownership	0.3	4.46	11.9	6.34	22.49
firms with state ownership	0.27	3.97	10.72	4.95	22.42
firms without foreign ownership	0.28	3.88	11.15	4.68	22.78
firms with foreign ownership	0.45	9.3	14.9	20.29	18.17
lower productivity firms	0.28	3.79	11.67	4.98	22.38
higher productivity firms	0.34	5.74	10.57	8.25	22.71

In Table 4.7 below there is the same descriptive statistics as in the table above, but only for the firms producing tradable goods: average level of diversity, localization and urbanization (average for the years 1999-2008).

Table 4.7. Average level of agglomeration for various types of firms; all tradables

	Mean (divpq)	**Mean (llocf)**	**Mean (llocnl)**	**Mean (lurbf)**	**Mean (lurbnl)**
'old' firms	0.33	4.98	11.41	6.57	22.81
'young' firms	0.30	4.38	10.76	6.35	22.30
firms without private ownership	0.33	5.54	12.10	6.90	23.01

	Mean (divpq)	Mean (llocf)	Mean (llocnl)	Mean (lurbf)	Mean (lurbnl)
firms with private ownership	0.31	4.49	10.85	6.34	22.50
firms without state ownership	0.31	4.43	10.82	6.45	22.36
firms with state ownership	0.33	5.31	11.75	6.55	23.07
firms without foreign ownership	0.31	4.40	10.93	5.29	23.12
firms with foreign ownership	0.45	8.91	13.98	20.06	17.09
lower productivity firms	0.30	4.33	11.42	5.70	22.61
higher productivity firms	0.36	6.21	10.36	9.12	22.68

In Table 4.6 all enterprises in the sample are considered, both those producing tradable goods and those producing non-tradable goods. 'Old' firms face slightly higher agglomeration levels (diversity, localization and urbanization) than 'young' firms in cities where they are located. As for the foreign firms, they are located in cities where agglomeration level is on average substantially higher, except for the indicator 'urbanization level of national firms'. Urbanization level of foreign firms is particularly high for firms with foreign ownership (20.29 vs. 4.68 for firms without foreign ownership), and localization level of foreign firms is rather high too (9.3 vs. 3.88 for firms without foreign ownership). Localization level of national firms is 14.9 for the firms with foreign ownership, while it is 11.15 for the firms without foreign ownership. Diversity for the firms with foreign ownership is 0.45, higher than for the other firms. Concerning productivity level, more productive firms are located in cities with slightly higher agglomeration, particularly agglomeration of foreign firms is higher; however, one indicator of agglomeration – localization level of the national firms – is lower.

Table 4.7 above shows that when only tradable industries are considered, diversity index is higher in most cases, and only in one case is equal to that for all enterprises (for the firms with foreign ownership). For localization indices there is no obvious pattern, if to compare enterprises belonging to tradable industries with overall sample of enterprises. Urbanization indices for enterprises belonging to tradable industries are higher than those for the overall sample of enterprises. Only the enterprises with foreign ownership belonging to tradable industries face slightly lower urbanization, and high productivity firms face very slightly lower urbanization arising from national firms. The patterns for different types of firms remain the same in most cases, except for comparison of the firms with private and state ownership.

Table 4.8. Correlation matrix of productivity and agglomeration variables; 'young' and 'old' firms

tradable==1 (obs=112374) – **all tradables**						
	productivity	divpq	llocf	llocnl	lurbf	lurbnl
productivity	1.0000					
divpq	0.0031	1.0000				
llocf	0.0265	0.7466	1.0000			
llocnl	0.0142	0.6613	0.5183	1.0000		
lurbf	0.0203	0.7560	0.9132	0.4982	1.0000	
lurbnl	0.0025	0.5856	0.3791	0.5227	0.2649	1.0000
new==1&tradable==1 (obs=37352) – **'young' enterprises**						
	productivity	divpq	llocf	llocnl	lurbf	lurbnl
productivity	1.0000					
divpq	0.0194	1.0000				
llocf	0.0321	0.6943	1.0000			
llocnl	0.0237	0.6035	0.4734	1.0000		
lurbf	0.0373	0.6999	0.8835	0.4402	1.0000	
lurbnl	-0.0003	0.5272	0.3101	0.4904	0.1614	1.0000
new==0&tradable==1 (obs=75022) - **'old' enterprises**						
	productivity	divpq	llocf	llocnl	lurbf	lurbnl
productivity	1.0000					
divpq	0.0005	1.0000				
llocf	0.0273	0.7697	1.0000			
llocnl	0.0133	0.6890	0.5395	1.0000		
lurbf	0.0185	0.7819	0.9270	0.5256	1.0000	
lurbnl	0.0038	0.6251	0.4225	0.5480	0.3283	1.0000

The correlations between productivity and agglomeration indices displayed in Table 4.8 above are higher for 'young' enterprises than for the 'old' enterprise. From the table above we also see relative importance of diversity and urbanization externalities for the 'young' enterprises, and relative importance of the localization externalities for the 'old' enterprises. The same time, these differences are not substantial. However, index of urbanization arising from concentration of national enterprises turned out to have negative correlation with productivity. It might be due to non-linear form of correlation. Otherwise, results for tradable industries are similar to those for all industries.

Table 4.9. Correlation matrix of productivity and agglomeration variables; firms with high productivity and firms with FDI

tradable==1 (obs=112374) – **all tradables**						
	productivity	divpq	llocf	llocnl	lurbf	lurbnl
productivity	1.0000					
divpq	0.0031	1.0000				
llocf	0.0265	0.7466	1.0000			
llocnl	0.0142	0.6613	0.5183	1.0000		
lurbf	0.0203	0.7560	0.9132	0.4982	1.0000	
lurbnl	0.0025	0.5856	0.3791	0.5227	0.2649	1.0000
highprod==1&tradable==1 (obs= 14264) **– firms with high productivity**						
	productivity	divpq	llocf	llocnl	lurbf	lurbnl
productivity	1.0000					
divpq	-0.0173	1.0000				
llocf	0.0352	0.7286	1.0000			
llocnl	0.0182	0.6193	0.5539	1.0000		
lurbf	0.0145	0.7340	0.8441	0.4917	1.0000	
lurbnl	0.0016	0.5132	0.3492	0.5596	0.1526	1.0000
fdi==1&tradable==1 (obs=8722) **- firms with FDI**						
	productivity	divpq	llocf	llocnl	lurbf	lurbnl
productivity	1.0000					
divpq	0.0298	1.0000				
llocf	0.0457	0.7131	1.0000			
llocnl	0.0281	0.6740	0.5914	1.0000		
lurbf	0.0329	0.7098	0.4783	0.5056	1.0000	
lurbnl	0.0203	0.7159	0.5676	0.9683	0.5412	1.0000

Table 4.9 shows that foreign firms have higher correlation of productivity with agglomeration indices than all firms belonging to tradable industries on average. As for more productive firms, correlations of productivity with localization indices for them are higher than for all firms on average, but with urbanization effects it is lower. The explanation may be that less productive firms are more dependent on the environment, and benefit greater from the favourable conditions created by presence of the other firms from the industries in the city, but suffer more from competition in the same industry.

4.4. Estimation issues: endogeneity

The origins of endogeneity, such as *unobserved heterogeneity* and *simultaneity*, were described in Chapters 2 and 3. Foreign firms have generally more possibility to choose their location. National firms do have certain flexibility as well in the market economy. Therefore we compared the distribution of 'young' and 'old' national firms; we found that they are very

similar (Chapter 3, Figure.3.2). However, distribution of national and foreign firms was found to vary (Chapter 3, Figure.3.3), therefore variables accounting for agglomeration effects generated by foreign firms are assumed to be endogenous.

Based on the relevance and validity conditions several instrumental variables were chosen. We tested the relevance of the instrumental variables by regressing each endogenous variable on a set of exogenous variables including the instrumental variables. Coefficients of at least several instrumental variables turned out to be significant regardless of choice of 1, 2, or 3 year lag. Coefficients were comparatively higher for 2 year lagged agglomeration variables, therefore they were chosen as the instrumental variables for agglomeration indices. Validity of the instruments was assumed to hold, because 2 year lagged agglomeration indices, localization quotient (LQ), Elison-Glaeser index, and the tourist potential variable are assumed to be uncorrelated with the error term.

The model is estimated using first differencing approach to deal with endogeneity arising from unobserved heterogeneity. Instruments with GMM method are used to deal with endogeneity arising from simultaneity. The results were compared with the approach using fixed effects and 2SLS, and first differencing with GMM was chosen.

4.5. Estimation results

In the analysis below urbanization coefficient will be used to reflect concentration of firms in cities, because it seems more suitable than diversity coefficient in situation when the number of national firms is much greater than that of foreign firms (the share of foreign firms in total number of firms is 0.08 – 0.09 in the sample with all firms producing tradable goods and in the sample with city level data). Firstly, the model is estimated using enterprise level fixed effects; the results are presented below.

Table 4.10. Regression results with home market potential (HMP) based on city level data; national and foreign firms

Dependent variable: ln(*revenue*)

	(1) All tradables	(2) Foreign firms	(3) National firms	(4) All tradables-1	(5) Foreign firms -1	(6) National firms -1
lfixed_assets	0.235***	0.294***	0.226***	0.233***	0.289***	0.225***
	(22.61)	(8.96)	(20.76)	(23.16)	(9.10)	(21.27)
llabour	0.428***	0.279***	0.457***	0.418***	0.275***	0.446***
	(32.03)	(8.56)	(31.20)	(32.12)	(8.75)	(31.21)
lhme	0.493***	0.218	0.504***	0.497***	0.412*	0.502***

	(1) All tradables	(2) Foreign firms	(3) National firms	(4) All tradables-1	(5) Foreign firms -1	(6) National firms -1
city						
	(9.30)	(0.75)	(9.39)	(10.07)	(1.71)	(9.98)
lwage city	-0.0405	-0.111	-0.0329	-0.0157	0.0318	-0.0205
	(-1.00)	(-0.64)	(-0.79)	(-0.42)	(0.19)	(-0.53)
core	9.600***	12.01***	9.309***	9.377***	9.878***	9.300***
	(15.35)	(7.14)	(13.90)	(16.79)	(7.21)	(15.41)
core2	-6.663***	-8.978***	-6.251***	-6.779***	-8.122***	-6.556***
	(-8.22)	(-5.38)	(-6.91)	(-8.93)	(-5.59)	(-7.79)
lloc	0.0322***	0.0503**	0.0291***			
	(4.57)	(2.36)	(3.87)			
lloc2	-0.00133***	-0.00241**	-0.00115***			
	(-3.54)	(-2.35)	(-2.83)			
lurbpq	0.280***	0.571***	0.268***			
	(10.64)	(3.78)	(10.09)			
llocf				0.00734**	0.0129**	0.00539
				(2.11)	(2.38)	(1.25)
llocnl				0.0154***	0.00710	0.0152***
				(2.74)	(0.22)	(2.64)
llocnl2				-0.000625*	-0.000875	-0.000546
				(-1.90)	(-0.52)	(-1.60)
lurbf				0.0463***	0.0278*	0.0604***
				(3.53)	(1.96)	(2.96)
lurbnl				0.228***	0.201**	0.236***
				(9.73)	(1.99)	(9.87)
a_road	0.000347***	0.000663**	0.000350***	0.000298***	0.000511*	0.000271**
	(3.44)	(2.13)	(3.25)	(2.92)	(1.71)	(2.39)
rw_road	-0.00214	-0.00920*	-0.00144	-0.00308**	-0.0126**	-0.00234*
	(-1.58)	(-1.72)	(-1.03)	(-2.31)	(-2.33)	(-1.70)
lbusn envrisk	-0.0109**	-0.0284	-0.00627	-0.0144***	-0.0394**	-0.00968*
	(-2.10)	(-1.57)	(-1.16)	(-2.83)	(-2.18)	(-1.81)
_cons	-5.858***	-3.378	-6.166***	-4.746***	2.462	-5.452***
	(-6.34)	(-0.81)	(-6.54)	(-5.51)	(0.59)	(-6.15)
firm fixed effects	yes	yes	yes	yes	yes	yes
N	49866	4215	45651	51922	4473	47449

	(1) All tradables	(2) Foreign firms	(3) National firms	(4) All tradables-1	(5) Foreign firms -1	(6) National firms -1
adj. R^2	0.428	0.385	0.436	0.435	0.390	0.444
sigma_u	1.151	1.759	1.097	1.451	2.671	1.305
sigma_e	0.522	0.610	0.511	0.516	0.597	0.507
rho	0.830	0.893	0.821	0.888	0.952	0.869

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

When localization externalities are considered as those generated by all firms (national and foreign) together and are reflected in the variables 'lloc' and 'lloc2', they are significant and have an inverted U shape (Models (1) – (3)). Urbanization externalities are significant and positive. Models (1) – (3) show that foreign firms benefit relatively more both from localization and from urbanization externalities. The reasons behind it might be that they have more choice where to locate, and that their level of human capital, particularly of management, might be higher (Vakhitov, 2008).

Concerning the models with externalities subdivided into those generated by national and by foreign firms (Models 4 - 6), urbanization effects generated by foreign firms and by national firms remain positive and significant; localization effects keep the inverted U shape only for the effects generated by the national firms in Model (4) for all tradables; in Model (5) for the foreign firms, localization effects generated by national firms are insignificant. In Model (6) for the national firms, localization effects generated by national firms are positive.

Localization externalities generated by foreign firms are significant and positive for the other foreign firms, i.e. firms with foreign ownership benefit from presence of foreign firms in the same industry. Localization effects generated by national firms are significant and positive for the other national firms. The reason behind it might be that national and foreign firms working in the same industry and the same city have to a certain extent different labour markets, different suppliers and do not interact intensively enough to share knowledge. Another reason for the insignificant effects might be that beside the potential positive effects generated by common suppliers, larger labour market and knowledge exchange, there is competition between national and foreign firms working in the same industry, particularly for resources, such as labour. While in the long run competition is assumed to stimulate improvements in production and management, in the short run it might negatively affect performance.

National and foreign firms benefit from urbanization externalities, from those generated both by national and foreign firms; for the foreign firms these effects are found to be weaker (it is a contradiction with Models (2) – (3), where for the foreign firms all effects

are found to be stronger). For the foreign firms urbanization effects generated by the national firms were found to be stronger than the urbanization effects generated by the other foreign firms. As urbanization variable is highly correlated with diversity variable, we may conclude that importance of urbanization economies reflects importance of the volume and diversity of economic activity in a city. In our sample it is economic activity in manufacturing industries.

HMP generated by the other cities is found to be positive and significant when all firms are considered together and for the national firms in all models; for the foreign firms this result is not robust (significant in Model (5), but not significant in Model (2)). HMP generated by the city itself is reflected in the agglomeration coefficient (localization and urbanization coefficients); the results are discussed above. Wage effect is found to be insignificant. As it was mentioned in Chapter 3, wage was chosen to be a proxy for human capital quality. However, the overall effect of wage on productivity is controversial. On one hand, higher wage means higher quality of human capital and might mean higher productivity. On the other hand, costs for enterprises arising from higher wage in a city might outweigh the first effect.

The impact of automobile roads is positive and significant as it was expected. The impact of railway roads turned out to be negative and insignificant. The risks of business environment are found to be significant in the model where all enterprises are considered together and in the Models (4) – (6), for all types of enterprises. The logic behind this indicator is that business owners and managers have to spend resources to deal with the difficulties of economic environment rather that to improve their products, production processes and management practices. For the national enterprises this indicator is much less significant, possibly because they are better adjusted to the economic realities and know how to circumvent the difficulties of business environment; however, the indicator remains significant for them too.

In Tables 4.11 **-** 4.13 below, it is analyzed, how agglomeration externalities generated by the national and foreign firms vary for different groups of industries (classification is presented in Chapter 2). In Table 4.11 both national and foreign firms are considered.

Table 4.11. Regression results with HMP based on city level data, by class of industry
Dependent variable: ln(*revenue*)

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
lfixed_ assets	0.233***	0.199***	0.233***	0.215***	0.226**	0.225***

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
	(23.16)	(8.34)	(11.45)	(6.45)	(2.59)	(14.83)
llabour	0.418***	0.493***	0.393***	0.407***	0.503***	0.435***
	(32.12)	(13.35)	(14.82)	(12.32)	(4.94)	(21.92)
lhme city	0.497***	0.368***	0.549***	0.433***	-0.917	0.410***
	(10.07)	(3.71)	(5.73)	(4.00)	(-1.07)	(4.15)
lwage city	-0.0157	0.148**	-0.0560	-0.000609	1.050*	0.100
	(-0.42)	(2.20)	(-0.78)	(-0.01)	(1.86)	(1.32)
core	9.377***	19.95***	11.22***	13.46***	9.467***	9.155***
	(16.79)	(7.73)	(9.66)	(6.00)	(4.84)	(10.47)
core2	-6.779***	-30.80***	-8.842***	-13.31***	-3.987*	-10.29***
	(-8.93)	(-6.06)	(-6.53)	(-4.00)	(-1.83)	(-8.72)
llocf	0.00734**	-0.00170	0.00464	0.0119	0.0668	0.0528**
	(2.11)	(-0.30)	(0.68)	(1.37)	(1.23)	(2.56)
llocf2					-0.00421	-0.00200***
					(-1.61)	(-2.79)
llocnl	0.0154***	0.0225	0.0347***	0.0385***	0.0962***	0.0153
	(2.74)	(1.44)	(3.60)	(2.70)	(2.97)	(0.97)
llocnl2	-0.000625*	-0.00126	-0.00200***	-0.00227***	-0.00611***	-0.000942
	(-1.90)	(-1.42)	(-3.60)	(-2.67)	(-3.38)	(-1.07)
lurbf	0.0463***	0.0643	0.0292*	0.0365*	-0.343	0.157***
	(3.53)	(1.29)	(1.87)	(1.82)	(-1.42)	(4.30)
lurbnl	0.228***	0.154***	0.277***	0.305***	0.348	0.249***
	(9.73)	(2.70)	(5.71)	(4.86)	(1.43)	(5.95)
a_road	0.000298***	0.000156	0.00000612	0.000497*	-0.000113	0.000172
	(2.92)	(0.54)	(0.03)	(1.93)	(-0.15)	(1.06)
rw_road	-0.00308**	-0.00369	-0.00272	0.00617*	0.00272	-0.00650***
	(-2.31)	(-1.08)	(-1.27)	(1.82)	(0.40)	(-2.95)
lbusn envrisk	-0.0144***	-0.0134	-0.0261**	-0.00892	-0.0799*	-0.0313***
	(-2.83)	(-0.93)	(-2.37)	(-0.77)	(-1.70)	(-4.11)
_cons	-4.746***	-1.256	-6.119***	-6.956***	17.26	-4.892***
	(-5.51)	(-0.69)	(-3.95)	(-3.64)	(1.44)	(-2.74)
firm fixed effects	yes	yes	yes	yes	yes	yes
N	51922	5786	13464	7072	828	17272

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
adj. R^2	0.435	0.458	0.470	0.466	0.452	0.443
sigma_u	1.451	1.251	1.734	1.963	4.289	1.923
sigma_e	0.516	0.473	0.493	0.492	0.593	0.528
rho	0.888	0.875	0.925	0.941	0.981	0.930

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

When groups of industries are treated separately, localization effects generated by foreign firms remain significant only for science based industries, with an inverted U shape. Localization effects generated by national firms turn out to be significant for basic materials, basic machinery, integrating machinery industries, but not for traditional and science based ones, and keep the inverted U shape. Urbanization externalities generated by foreign firms proved to be significant for basic materials, basic machinery, science based industries. As for urbanization externalities generated by national firms, they are found to be significant for traditional, basic materials, basic machinery, science based industries, i.e. for all industries except for integrating machinery. Urbanization externalities turned out to be important for various types of industries. Lack of observations for integrating machinery might distort the results. HMP arising from the other cities proved to be the largest for the basic materials industries, followed by basic machinery, science based, and traditional industries.

According to the results presented in Table 4.12 below, national firms seem to benefit relatively more from interaction with the other national firms than with the foreign firms, which is in line with the overall results discussed earlier. The results suggest that it might be true not only for localization externalities, but for the urbanization externalities as well. National firms in basic materials, basic machinery and integrating machinery are found to benefit from localization economies generated by the other national firms; national firms in traditional, basic materials, basic machinery and science based industries benefit from urbanization externalities generated by the other national firms.

However, national firms do benefit from externalities generated by the foreign enterprises as well. Localization externalities generated by foreign firms proved to be beneficial for the national firms in science based industries; urbanization economies generated by the foreign firms proved to be positive and significant for the national firms belonging to science based and traditional industries. Urbanization externalities reflect that both national and foreign firms benefit from large volume of activities in a city, highly correlated with diversity of activities leading to exchange of ideas etc. HMP arising from the other cities

proved to be the largest for the basic materials industries, followed by science based, basic machinery and traditional industries.

Table 4.12. Regression results with HMP based on city data, by class of industry, national firms

Dependent variable: ln(*revenue*)

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
lfixed_assets	0.225***	0.176***	0.227***	0.211***	0.223**	0.219***
	(21.27)	(7.38)	(10.22)	(6.06)	(2.20)	(13.93)
llabour	0.446***	0.505***	0.418***	0.448***	0.426***	0.466***
	(31.21)	(12.52)	(14.42)	(12.05)	(5.36)	(21.96)
lhme city	0.502***	0.379***	0.524***	0.449***	-0.126	0.452***
	(9.98)	(3.75)	(5.29)	(4.28)	(-0.27)	(4.52)
lwage city	-0.0205	0.140**	-0.0610	0.00648	0.559	0.0646
	(-0.53)	(2.14)	(-0.81)	(0.09)	(1.52)	(0.84)
core	9.300***	19.02***	11.22***	13.59***	7.823***	9.392***
	(15.41)	(7.16)	(8.89)	(5.85)	(3.60)	(10.56)
core2	-6.556***	-29.03***	-8.597***	-13.48***	-2.500	-10.64***
	(-7.79)	(-4.92)	(-5.72)	(-3.96)	(-1.17)	(-8.79)
llocf	0.00539	-0.00578	0.00489	0.00729	0.0614	0.0476*
	(1.25)	(-0.96)	(0.46)	(0.75)	(0.66)	(1.76)
llocf2					-0.00434	-0.00189**
					(-1.01)	(-2.14)
llocnl	0.0152***	0.0198	0.0339***	0.0409***	0.0941***	0.0118
	(2.64)	(1.27)	(3.47)	(2.83)	(2.94)	(0.75)
llocnl2	-0.000546	-0.000978	-0.00193***	-0.00235***	-0.00540***	-0.000775
	(-1.60)	(-1.12)	(-3.36)	(-2.71)	(-3.02)	(-0.87)
lurbf	0.0604***	0.145***	0.0381	0.0349	-0.147	0.174***
	(2.96)	(3.75)	(0.79)	(1.41)	(-0.80)	(4.55)
lurbnl	0.236***	0.166***	0.310***	0.292***	0.167	0.252***
	(9.87)	(2.81)	(5.99)	(4.45)	(1.33)	(5.91)
a_road	0.000271**	0.0000275	-0.0000144	0.000471*	-0.000438	0.000137
	(2.39)	(0.10)	(-0.06)	(1.73)	(-0.58)	(0.83)
rw_road	-0.00234*	-0.00304	-0.00244	0.00811**	0.00257	-0.00571**
	(-1.70)	(-0.89)	(-1.07)	(2.41)	(0.37)	(-2.56)

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
lbusn envrisk	-0.00968*	-0.00600	-0.0218*	-0.00884	-0.0817*	-0.0230***
	(-1.81)	(-0.42)	(-1.82)	(-0.75)	(-1.66)	(-3.00)
_cons	-5.452***	-2.144	-6.744***	-7.791***	6.574	-6.088***
	(-6.15)	(-1.24)	(-4.16)	(-4.24)	(0.83)	(-3.37)
firm fixed effects	yes	yes	yes	yes	yes	yes
N	47449	5258	12011	6492	739	16240
adj. R^2	0.444	0.476	0.476	0.478	0.432	0.456
sigma_u	1.305	1.590	1.256	1.892	1.894	2.089
sigma_e	0.507	0.447	0.490	0.487	0.516	0.516
rho	0.869	0.927	0.868	0.938	0.931	0.942

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Concerning Table 4.13 below, the number of observations is not so large for the foreign firms, particularly when separate industries are considered. The results for integrating machinery are disputable in Tables 4.11 - 4.13 due to lack of observations; in Table 4.13 there is an evident problem with them so they are not considered. When industries are considered separately, localization externalities generated by foreign firms are significant for the other foreign firms belonging to science based industries; they have and inverted U shape. Urbanization externalities generated by foreign firms are significant for basic machinery. Localization externalities generated by national firms are not significant, while urbanization externalities generated by national firms are significant for basic machinery and science based industries. HMP arising from the other cities is positive and significant for basic materials industries, however, for the other industries it turned out to be either insignificant, or significant and negative. The reason behind low significance of various coefficients might be lack of observations for the foreign firms.

Table 4.13. Regression results with HMP based on city data, by class of industry, foreign firms

Dependent variable: ln(*revenue*)

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
lfixed_assets	0.289***	0.413***	0.269***	0.192***	-0.0982	0.247***
	(9.10)	(5.78)	(6.89)	(3.64)	(-0.56)	(5.13)

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
llabour	0.275***	0.400***	0.276***	0.171***	0.619**	0.262***
	(8.75)	(4.71)	(4.17)	(4.09)	(2.90)	(5.25)
lhme city	0.412*	0.197	0.630*	0.437	-9.127*	-1.184*
	(1.71)	(0.26)	(1.80)	(0.77)	(-1.81)	(-1.82)
lwage city	0.0318	0.131	-0.0510	-0.112	5.292	1.425***
	(0.19)	(0.20)	(-0.19)	(-0.21)	(1.70)	(2.89)
core	9.878***	23.27***	10.98***	18.51***	36.59***	7.374
	(7.21)	(4.05)	(4.80)	(4.31)	(5.07)	(1.37)
core2	-8.122***	-36.04***	-9.763***	-37.05***	-55.47***	-7.554
	(-5.59)	(-4.25)	(-4.56)	(-3.66)	(-3.11)	(-1.25)
llocf	0.0129**	0.0417	0.00358	0.0116	0.0504	0.0624**
	(2.38)	(1.52)	(0.71)	(0.61)	(1.03)	(2.43)
llocf2					0.000310	-0.00260*
					(0.07)	(-1.69)
llocnl	0.00710	0.0507	0.0111	-0.0876	7.369***	0.160
	(0.22)	(0.86)	(0.18)	(-1.65)	(3.25)	(1.34)
llocnl2	-0.000875	-0.00371	-0.00120	0.00340	-0.195***	-0.00929
	(-0.52)	(-1.02)	(-0.39)	(1.07)	(-3.40)	(-1.35)
lurbf	0.0278*	0.0184	0.0269	0.0199**	-0.655	-0.0355
	(1.96)	(0.41)	(1.55)	(2.50)	(-1.32)	(-0.22)
lurbnl	0.201**	-0.0554	0.136	0.443**	4.016**	0.481**
	(1.99)	(-0.16)	(0.94)	(2.26)	(2.40)	(2.40)
a_road	0.000511*	0.00151	0.0000907	0.000550	-0.000784	0.000738
	(1.71)	(1.29)	(0.20)	(0.69)	(-0.46)	(1.05)
rw_road	-0.0126**	-0.0157	-0.00329	-0.0205***	0.0694	-0.0302***
	(-2.33)	(-0.88)	(-0.60)	(-3.18)	(0.71)	(-3.96)
lbusn envrisk	-0.0394**	-0.0468	-0.0259	0.0183	0.0418	-0.141***
	(-2.18)	(-0.82)	(-1.05)	(0.35)	(0.24)	(-3.61)
_cons	2.462	9.550	-2.900	3.149	34.13	29.73**
	(0.59)	(0.58)	(-0.48)	(0.32)	(0.54)	(2.36)
firm fixed effects	yes	yes	yes	yes	yes	yes
N	4473	528	1453	580	89	1032
adj. R^2	0.390	0.416	0.441	0.424	0.690	0.344

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
sigma_u	2.671	3.679	1.724	4.983	78.34	4.280
sigma_e	0.597	0.650	0.517	0.513	0.816	0.663
rho	0.952	0.970	0.918	0.989	1.000	0.977

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In the tables above, enterprise fixed effects were applied to deal with endogeneity arising from unobserved heterogeneity. Below we take into account also endogeneity arising from simultaneity by using instrumental variables, and compare 2SLS and GMM methods.

Table 4.14. Regression results with firm fixed effects and instruments
Dependent variable: ln(*revenue*)

	(1) OLS	(2) 2SLS (f.e.,IVs: llocnl2lag2 LQ EGcities)	*(3) GMM (f.e.,IVs: llocnl2lag2 LQ EGcities)*	(4) 2SLS (f.e.,IVs: llocnl2lag2 lurbnllag2 EGcities)	(5) GMM (f.e.,IVs: llocnl2lag2 lurbnllag2 EGcities)
core	9.377***	11.51***	*13.64****	9.239***	8.957***
	(16.79)	(6.06)	*(8.93)*	(4.20)	(4.09)
core2	-6.779***	-8.235***	*-10.15****	-6.114***	-5.893***
	(-8.93)	(-4.50)	*(-6.66)*	(-2.86)	(-2.76)
llocnl	0.0154***	0.0980***	*0.0982****	0.0997***	0.0988***
	(2.74)	(4.60)	*(4.61)*	(4.50)	(4.46)
llocnl2	-0.000625*	-0.00643***	*-0.00641****	-0.00662***	-0.00656***
	(-1.90)	(-4.26)	*(-4.25)*	(-4.23)	(-4.19)
lurbnl	0.228***	0.322***	*0.358****	0.313***	0.309***
	(9.73)	(6.95)	*(8.53)*	(7.52)	(7.44)
llocf	0.00734**	0.226	*-0.108*	0.600	0.646
	(2.11)	(0.66)	*(-0.37)*	(1.42)	(1.53)
lurbf	0.0463***	1.622***	*2.033****	1.235***	1.149***
	(3.53)	(3.95)	*(5.84)*	(3.42)	(3.23)
lfixed_assets	0.233***	0.214***	*0.216****	0.211***	0.212***
	(23.16)	(16.55)	*(16.69)*	(15.90)	(15.96)
llabour	0.418***	0.421***	*0.425****	0.417***	0.415***
	(32.12)	(27.06)	*(27.72)*	(27.62)	(27.61)
lhmecity	0.497***	0.669***	*0.657****	0.665***	0.671***
	(10.07)	(7.02)	*(6.90)*	(7.31)	(7.38)

	(1) OLS	(2) 2SLS (f.e.,IVs: llocnl2lag2 LQ EGcities)	*(3) GMM (f.e.,IVs: llocnl2lag2 LQ EGcities)*	(4) 2SLS (f.e.,IVs: llocnl2lag2 lurbnllag2 EGcities)	(5) GMM (f.e.,IVs: llocnl2lag2 lurbnllag2 EGcities)
lwagecity	-0.0157	-0.400***	*0.657****	-0.357***	-0.354***
	(-0.42)	(-3.45)	*(6.90)*	(-3.42)	(-3.38)
a_road	0.000298***	-0.00400***	*-0.00459****	-0.00349***	-0.00332***
	(2.92)	(-4.16)	*(-5.05)*	(-4.25)	(-4.10)
rw_road	-0.00308**	-0.0172***	*-0.0200****	-0.0142***	-0.0135***
	(-2.31)	(-4.74)	*(-6.02)*	(-4.27)	(-4.12)
lbusnenv risk	-0.0144***	-0.121***	*-0.136****	-0.108***	-0.103***
	(-2.83)	(-4.97)	*(-5.84)*	(-5.10)	(-4.94)
_cons	-4.746***				
	(-5.51)				
firm fixed effects	yes	yes	*yes*	yes	yes
N	51922	51812	*51812*	51812	51812
adj. R^2	0.435	-1.066	*-1.705*	-1.199	-1.223
sigma_u	1.451				
sigma_e	0.516	0.902	*1.032*	0.931	0.936
rho	0.888				
Kleibergen-Paap underidentification test (p-value)		22.114 (0.0000)	*22.114 (0.0000)*	13.781 (0.0010)	13.781 (0.0010)
Kleibergen-Paap weak identification test		6.668	*6.668*	4.595	4.595
Hansen overidentification test (p-value)		3.562 (0.0591)	*3.562 (0.0591)*	1.776 (0.1826)	1.776 (0.1826)

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Regression results show that OLS and IV estimates differ, the difference being especially great for the foreign firms' agglomeration variables llocf and lurbf. Hausman specification test also rejects the null hypothesis of negligible difference between the coefficients of OLS and instrumental variable estimation (2SLS or GMM). It confirms the endogeneity in the model leading to the necessity of using the instrumental variables.

Although OLS estimates would be consistent and more efficient if the null hypothesis were confirmed, in case of rejecting the null hypothesis they are inconsistent, while 2SLS-IV or GMM-IV methods give consistent estimates.

Firstly, we consider Model (2), first stage estimation with dependent variables llocf and lurbf that are assumed to be endogenous. F test of excluded instruments and Angrist-Pischke multivariate F test of excluded instruments show that both endogenous variables, llocf and lurbf, are identified, as for each of these regressors the null hypothesis that the endogenous regressor is unidentified is rejected.

Kleibergen-Paap Wald rk F statistic of 6.67 exceeds the Stock-Yogo weak ID test critical value of 5.45 for 25% maximal IV size and of 6.40 for 20% maximal IV size. As weak identification should not be a problem here, then bias of the coefficients of llocf and lurbf is not very likely to occur. From the 2nd stage estimation results we see that llocf is not significant, while 1% increase of the foreign firms urbanization level leads to 1.62% average increase in revenue of all firms in a city. In other words, if the foreign firms' urbanization level increases twice (100%), then revenue increases 162% on average. The same time, 1% increase in national firms' urbanization level leads to 0.322% increase in revenue, i.e. if national firms' urbanization level increases twice, revenue increases 32.2%. Compared to OLS estimates, the 2SLS-IV estimates give much higher coefficient of lurbf and slightly higher coefficient of lurbnl (OLS results: Model (1)).

High significance of the foreign firms' urbanization level can be explained by the fact that existence of foreign firms is favourable for development of suppliers for enterprises belonging to various industries working in a city, such as services for business. Besides, by bringing in new business practices they contribute to formation of favorable business climate. Referring to the theoretical background of FDI spillovers, the effects generated by the foreign firms' urbanization level can be associated partly with vertical spillovers and supply-backward spillovers between suppliers and final producers (Merlevede, Schoors, 2008). In this context, localization externalities can be associated with horizontal spillovers from FDI, described in more detail in literature review (Chapter 1).

A possible reason for insignificance of llocf is that national and foreign firms in the same industry do not interact closely enough. Another possible reason of the insignificant result is that there are both positive and negative effects for the national firms generated by the foreign firms' presence, and these effects compensate each other. Positive effects can be associated with knowledge spillovers and competition leading to enhanced efficiency, negative effects being costs of competition for resources and for consumers. The same time, the OLS estimates (here inconsistent) provide positive though not highly significant

coefficient of llocf. Concerning the llocnl estimates, they are higher with 2SLS and GMM IV estimation than with OLS estimation. For instance, 2SLS estimation (Model (2)) shows that 1% increase in llocnl lead to 0.098% increase in revenue on average for all enterprises, while for OLS the coefficient is 0.0154. The effects in both cases have an inverted U shape. Tests of joint significance of endogenous regressors in the main equation reject the null hypothesis that the regressors are insignificant.

Concerning the second stage, estimation by 2SLS (Table 4.14, Models (2) and (4)) and GMM (Table 4.14**,** Models (3) and (5)) methods give similar results, as well as using different sets of instruments (llocnl2lag2 LQ EGcities in Models (2) and (3); llocnl2lag2 lurbnllag2 EGcities in Models (4) and (5)). Underidentification test provides Kleibergen-Paap rk LM statistic equal to 22.114 with chi-sq(2) p-value equal to zero, rejecting the null hypothesis that the equation is underidentified. In other words, the equation is identified, i.e., the instrumental variables are correlated with the endogenous regressors. Weak identification test provides Kleibergen-Paap rk Wald F statistic equal to 6.668, which exceeds the Stock-Yogo weak ID test critical value of 6.40 for 20% maximal IV size, showing that there is no problem of weak identification in the model.

Hansen J statistic for overidentification test of all instruments is equal to 3.562 with chi-sq(1) p-value equal to 0.0591. The joint null hypothesis is that the instruments are valid (uncorrelated with the error term) and that the excluded instruments are correctly excluded from the equation. Here the null hypothesis can be accepted on 95% confidence interval (as p-value is greater than 0.05). In Model (3), GMM-IV estimation with the same instrumental variables gives similar results, and the outcomes of the tests are similar.

As p-value for Hansen J statistic is rather low, we compare the specification discussed above with the model where the instruments llocnl2lag2 lurbnllag2 EGcities are used (Models (4), (5)) in order verify reliability of the instruments. For 2SLS with these instrumental variables, instruments are found to be weak (Weak identification test: Kleibergen-Paap rk Wald F statistic is equal to 4.595, which is less than the lowest Stock-Yogo weak ID test critical value for 25% maximal IV size equal to 5.45). However, Hansen J statistic shows that the instruments are valid and uncorrelated with the error term. In this model Hansen J statistic is 1.776 and p value is 0.1826; therefore the hypothesis is accepted with less risk of making an error than in the models 2 and 3. The coefficients in these models are closer to the OLS estimates; there might be a downward bias (Dickson, 2009, p. 28).

Overall, the results of the Models (1) – (5) are robust in terms of signs and magnitude of the coefficients, making it possible to draw conclusions, though cautiously, about the effects generated by economic agglomeration in the cities. In Table 4.15 below, first

differencing is applied; GMM is used; the variables llocf and lurbf are instrumented; the instruments are llocnl2lag2 llocnllag2 LQ lurbflag2. Models (3) and (4) are estimated with time fixed effects.

Table 4.15. Regression results with first differencing and instruments
Dependent variable: D. ln(*revenue*)

	(1) All tradables	(2) All tradables (without Moscow)	(3) All tradables	(4) All tradables (without Moscow)	(5) National firms	(6) Foreign firms
D.llocf	-0.674***	-0.504**	-0.750***	-0.554*	-0.731***	-0.220
	(-2.68)	(-1.98)	(-2.80)	(-1.94)	(-2.62)	(-1.07)
D.lurbf	1.360***	1.410***	1.362**	0.380	0.613***	-0.0291
	(2.69)	(3.03)	(2.34)	(0.59)	(2.73)	(-0.29)
D.lfixed_ assets	0.188***	0.161***	0.185***	0.164***	0.182***	0.210***
	(14.61)	(13.62)	(14.17)	(12.89)	(14.25)	(5.23)
D.llabour	0.245***	0.263***	0.249***	0.271***	0.268***	0.159***
	(15.72)	(19.45)	(15.82)	(19.81)	(19.87)	(5.52)
D.lhme city	0.299*	0.408***	0.390*	0.105	0.254*	-1.172
	(1.83)	(3.14)	(1.77)	(0.83)	(1.90)	(-1.33)
D.lwage city	-0.144	-0.00875	-0.126	0.0203	0.0684	0.121
	(-1.19)	(-0.35)	(-0.99)	(0.55)	(1.37)	(0.73)
D.core	14.90***	13.26***	15.12***	13.13***	13.83***	11.79***
	(8.06)	(10.17)	(7.79)	(9.47)	(8.06)	(4.39)
D.core2	-10.93***	-9.519***	-11.07***	-9.642***	-10.75***	-9.097***
	(-7.03)	(-8.59)	(-6.82)	(-8.00)	(-6.04)	(-5.25)
D.llocnl	0.0314***	0.0258***	0.0321***	0.0214***	0.0205***	0.0483
	(3.73)	(5.51)	(3.49)	(3.70)	(3.26)	(1.10)
D.llocnl2	-0.00237***	-0.00196***	-0.0024***	-0.00158***	-0.00146***	-0.00294
	(-3.51)	(-5.43)	(-3.40)	(-3.61)	(-3.31)	(-1.27)
D.lurbnl	0.358***	0.300***	0.320***	0.278***	0.267***	0.244*
	(7.12)	(8.79)	(5.51)	(7.87)	(7.31)	(1.82)
D.a_road	-0.000806	0.000235	-0.000892	0.0000637	0.000300	0.000142
	(-1.58)	(1.51)	(-1.20)	(0.27)	(1.03)	(0.27)
D.rw_ road	-0.0108***	-0.00809***	-0.00877***	-0.00506**	-0.00465*	-0.0204**
	(-3.61)	(-4.32)	(-2.60)	(-2.14)	(-1.91)	(-2.42)

D.lbusn envrisk	-0.0320***	-0.0335***	-0.0278*	-0.0218*	-0.00101	-0.0129
	(-3.04)	(-3.26)	(-1.76)	(-1.75)	(-0.13)	(-0.63)
_cons	-0.00493	-0.0313	-0.0818**	-0.0856***	-0.102***	0.0782
	(-0.14)	(-1.27)	(-2.29)	(-2.78)	(-3.57)	(0.59)
time fixed effects	no	no	yes	yes	yes	yes
N	41349	32720	41349	32720	37830	3519
adj. R^2	-2.423	-2.152	-2.775	-0.837	-1.340	-0.323
sigma_u						
sigma_e	1.030	0.953	1.082	0.728	0.842	0.718
rho						
Kleibergen-Paap underidentification test (p-value)	15.447 (0.0015)	5.412 (0.0668)	14.447 (0.0024)	19.820 (0.0002)	27.940 (0.0000)	5.068 (0.1669)
Kleibergen-Paap weak identification test	2.032	1.807	2.005	4.491	5.247	1.224
Hansen overidentification test (p-value)	0.369 (0.8313)	4.361 (0.0368)	0.002 (0.9990)	3.380 (0.1845)	5.042 (0.0804)	3.635 (0.1625)

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In the models considered below, the following variables were used as instruments: llocnl2lag2, llocnllag2, lurbflag2, LQ; the models were estimated with GMM. Model (1) was estimated for all enterprises producing tradable goods. 1st stage. F test of excluded instruments and Angrist-Pischke multivariate F test of excluded instruments show that both endogenous variables, llocf and lurbf, are identified, as for each of them the null hypothesis that the endogenous regressor is unidentified is strongly rejected. 2nd stage. Kleibergen-Paap underidentification test rejects the null hypothesis that the model is underidentified. Kleibergen-Paap Wald rk F statistic of 2.032 for weak identification test does not exceed the lowest Stock-Yogo weak ID test critical value of 4.73 that implies 30% maximal IV relative bias. Therefore coefficients may be biased, and the results should be treated cautiously. Hansen J statistic shows that the instruments are valid and uncorrelated with the error term.

Model (2) was estimated for enterprises producing tradable goods, excluding enterprises located in Moscow. 1st stage. For localization coefficient the hypothesis that the endogenous regressor is unidentified is strongly rejected both by F test of excluded

instruments and Angrist-Pischke multivariate F test of excluded instruments. However, for urbanization coefficient, F test of excluded instruments rejects this hypothesis only at 90% confidence interval. Angrist-Pischke multivariate F test of excluded instruments does not reject the hypothesis that the endogenous regressor is unidentified**.** 2nd stage. Kleibergen-Paap underidentification test rejects the null hypothesis that the model is underidentified on 90% confidence interval. Kleibergen-Paap Wald rk F statistic of 1.807 for weak identification test does not exceed the lowest Stock-Yogo weak ID test critical value of 5.45 that implies 25% maximal IV size. Therefore coefficients may be biased, and the results should be treated cautiously. Hansen J statistic 4.361 with p-value of 0.0368 shows that the joint null hypothesis that the instruments are valid instruments, can be accepted on 99% confidence interval, though it can be rejected on 95% confidence interval, therefore we cannot be confident that instruments are valid and uncorrelated with the error term.

In Models (3) - (6), time fixed effects are taken into account; the sample contains data for the years 2002-2008. 1st stage. For both models F test of excluded instruments and Angrist-Pischke multivariate F test of excluded instruments show that both endogenous variables, llocf and lurbf, are identified, as for each of them the null hypothesis that the endogenous regressor is unidentified is strongly rejected. At the 2nd stage, for Models (3) and (4), Kleibergen-Paap underidentification test rejects the null hypothesis that the model is underidentified. For Model (3), Kleibergen-Paap Wald rk F statistic of 2.005 for weak identification test does not exceed the lowest Stock-Yogo weak ID test critical value of 4.73 for 30% maximal IV relative bias. Therefore coefficients may be biased, and the results should be treated cautiously. For Model (4), Kleibergen-Paap Wald rk F statistic of 4.491 does not exceed the lowest Stock-Yogo weak ID test critical value of 4.73 for 30% maximal IV relative bias. Therefore coefficients may be biased, and the results should be treated cautiously. For Models (3) and (4), Hansen J statistic shows that the instruments are valid and uncorrelated with the error term.

Concerning the instruments, tourist potential turned out to be a suitable instrument for lurbf. However, if tourist potential is used, for example, instead of LQ, then llocf is unidentified, and the results are not valid. Using LQ is disputable; however, the tests described above showed that it is a valid instruments.

Model (5) is estimated for the national firms; the sample includes Moscow. At the 1st stage, for both models F test of excluded instruments and Angrist-Pischke multivariate F test of excluded instruments show that both llocf and lurbf are identified, as for each of these regressors the null hypothesis that the endogenous regressor is unidentified is strongly rejected. At the 2nd stage, Kleibergen-Paap underidentification test rejects the null hypothesis

that the model is underidentified. Kleibergen-Paap Wald rk F statistic of 5.247 for weak identification test exceeds the lowest Stock-Yogo weak ID test critical value of 4.73 that implies 30% maximal IV relative bias. Hansen J statistic 5.042 with p-value of 0.0804 shows that the joint null hypothesis that the instruments are valid instruments and uncorrelated with the error term, can be accepted on 95% confidence interval, though it can be rejected on 90% confidence interval. Overall, the tests show that the quality of the model is good.

For the national firms, 1% increase in presence of foreign firms in the same industry in a city (localization) leads to 0.731% decrease in revenue. However, 1% increase in presence of foreign firms in different industries in a city leads to 0.613% increase in revenue. Both coefficients are significant on 99% confidence interval. As for HMP, its 1% increase leads to 0.25% increase in revenue, on 90% confidence interval. Increase in a share of an industry *j* in the total revenue in a city *z* (the variable 'core') by 1% leads to increase in revenue by 0.138% up to a point when this share is 13.83/(2*10.75)=0.64; further concentration by 1% leads to decrease in positive effect by 0.107%, i.e. the effect has inverted U shape.

Localization of national firms also has inverted U shape. 1% increase in national firms' localization leads to 0.02% increase in revenue, optimum is at 0.0205/(2*0.00146)=7.02, i.e. exp(7.02)=1,119.4 roubles, which is an unrealistically low value, after reaching which, positive effects of localization start to diminish. 1% increase in national firms' presence in different industries in a city (urbanization) leads to 0.267% increase in revenue of the other national firms, the coefficient is significant on 99% confidence interval. The effect is lower than that from foreign firms' presence (0.613% increase).

For the foreign firms (Model (6)) tests show that llocf and lurbf are underidentified, that instruments are weak. However, Hansen J statistic 3.635 with p-value 0.1625 shows that the joint null hypothesis that the instruments are valid instruments and uncorrelated with the error term, can be accepted on 90% confidence interval. Therefore the results are biased and should be treated cautiously. However, the results remain robust to changes in instrumental variables. They show that for the foreign firms, a share of an industry *j* in the total revenue in a city *z* (the variable 'core') is significant, the effect having an inverted U shape. Urbanization arising from national firms (the variable 'lurbnl') turned out to be significant too, on 90% confidence interval, with positive sign. Compared to the model estimated with fixed effects and without instruments (Table 4.10 above), the result of 'core' is only slightly higher; the result of lurbnl is almost the same; it was only slightly lower in the model estimated with fixed effects and without instruments: coefficient 0.201 significant on 95% confidence interval, whereas here the coefficient is 0.244 significant on 90% confidence interval. In

Model (6) the other agglomeration variables and HMP are not significant; in the model estimated with fixed effects and without instruments, HMP and foreign firms' localization and urbanization, were positive and significant, on 90% and 95% confidence intervals. Business environment risk is significant and has negative effect in Models (1) – (4), i.e. when national and foreign firms are considered together.

The next questions are, do the effects from the foreign firms' presence vary across cities, and if yes, in which types of cities do foreign firms affect the most all firms producing tradable goods, and particularly national firms? The results for the city types are presented in table Table 4.16 below; national firms across the types of cities are considered in Table 4.17**.**

Table 4.16. Regression results for different types of cities; all tradables
Dependent variable: D. ln(*revenue*)

	(1) All tradables	(2) Agglomeration centers	(3) Cities within agglomerations, not centers	(4) Monotowns	(5) Other cities
D.llocf	-0.750***	0.899**	-0.663	-0.176	-0.483
	(-2.80)	(2.38)	(-0.39)	(-0.65)	(-1.11)
D.lurbf	1.362**	-0.740	-0.825	-0.336	-0.163
	(2.34)	(-1.06)	(-0.44)	(-0.78)	(-0.56)
D.lfixed_ assets	0.185***	0.198***	0.226***	0.233***	0.163***
	(14.17)	(7.97)	(4.37)	(6.81)	(9.40)
D.llabour	0.249***	0.207***	0.147*	0.218***	0.314***
	(15.82)	(8.14)	(1.87)	(6.17)	(16.00)
D.lhmecity	0.390*	-0.475	0.933	0.621**	0.262
	(1.77)	(-1.01)	(0.70)	(1.97)	(1.32)
D.lwagecity	-0.126	-0.775	-0.0383	0.0114	0.133
	(-0.99)	(-0.72)	(-0.39)	(0.06)	(1.08)
D.core	15.12***	-14.08	11.29***	5.443***	14.21***
	(7.79)	(-1.47)	(2.98)	(2.73)	(8.47)
D.core2	-11.07***	24.74	-11.02**	-1.764	-12.28***
	(-6.82)	(1.60)	(-1.99)	(-1.56)	(-6.36)
D.llocnl	0.0321***	0.0272	0.0462**	0.0183**	0.0190***
	(3.49)	(0.79)	(2.51)	(2.02)	(2.71)
D.llocnl2	-0.00242***	-0.00134	-0.00415***	-0.000751	-0.00136***
	(-3.40)	(-0.58)	(-3.19)	(-1.06)	(-2.65)
D.lurbnl	0.320***	-0.184	0.265**	0.224**	0.257***
	(5.51)	(-0.85)	(2.39)	(2.15)	(6.74)

	(1) All tradables	(2) Agglomeration centers	(3) Cities within agglomerations, not centers	(4) Monotowns	(5) Other cities
D.a_road	-0.000892	-0.000244	-0.00157	-0.000710	0.000111
	(-1.20)	(-0.18)	(-0.34)	(-1.22)	(0.30)
D.rw_road	-0.00877***	0.0320	-0.0288	-0.00186	-0.00451**
	(-2.60)	(1.63)	(-0.42)	(-0.41)	(-2.08)
D.lbusnenv risk	-0.0278*	0.0265	0.0450	-0.0288	-0.0148
	(-1.76)	(1.36)	(0.67)	(-1.39)	(-1.25)
_cons	-0.0818**	0.334	-0.0571	-0.158**	-0.160**
	(-2.29)	(1.48)	(-0.19)	(-2.39)	(-2.53)
time fixed effects	yes	yes	yes	yes	yes
N	41349	19698	1687	2294	18042
adj. R^2	-2.775	-4.079	-0.662	0.205	-0.142
sigma_u					
sigma_e	1.082	1.257	0.653	0.429	0.607
rho					
Kleibergen-Paap underidentification test (p-value)	14.447 (0.0024)	35.388 (0.0000)	3.141 (0.3704)	7.107 (0.0686)	5.140 (0.1618)
Kleibergen-Paap weak identification test	2.005	4.134	0.743	1.471	1.090
Hansen overidentification test (p-value)	0.002 (0.9990)	4.458 (0.1076)	2.129 (0.3448)	0.607 (0.7382)	0.900 (0.6378)

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Model (1) was commented in Table 4.15 above. The effects for the enterprises located in the agglomeration centers are reflected in Model (2). On the 1st stage, F test of excluded instruments and Angrist-Pischke multivariate F test of excluded instruments show that both llocf and lurbf are identified, as for each of these regressors the null hypothesis that the endogenous regressor is unidentified is strongly rejected. At the 2nd stage, Kleibergen-Paap underidentification test rejects the null hypothesis that the model is underidentified. Kleibergen-Paap Wald rk F statistic of 4.134 for weak identification test does not exceed the lowest Stock-Yogo weak ID test critical value of 4.73 for 30% maximal IV relative bias. Hansen J statistic 4.458 with p-value of 0.1076 shows that the joint null hypothesis that the

instruments are valid instruments and uncorrelated with the error term can be accepted on 90% confidence interval.

Localization effect generated by the foreign firms turned out to be positive and significant on 95% confidence interval, implying that 1% increase in foreign firms' presence in the same industry leads to 0.899% increase in revenue of national and foreign firms. When all cities are considered together, this effect is negative as it was mentioned above. Other agglomeration coefficients turned out to be insignificant.

The effects for the enterprises located in the cities within agglomerations but outside agglomeration centers are studied in Model (3). At the 1st stage, F test of excluded instruments and Angrist-Pischke multivariate F test of excluded instruments show that the hypothesis that llocf is unidentified cannot be rejected. According to F test of excluded instruments, the null hypothesis that lurbf is underidentified can be accepted, while with Angrist-Pischke multivariate F test it can be rejected on 95% confidence interval. At the 2nd stage, Kleibergen-Paap underidentification test cannot reject the null hypothesis that the model is underidentified. Kleibergen-Paap Wald rk F statistic of 0.743 for weak identification test does not exceed the lowest Stock-Yogo weak ID test critical value of 4.73 for 30% maximal IV relative bias. Therefore, the results are probably biased. However, Hansen J statistic 2.129 with p-value of 0.3448 shows that the joint null hypothesis that the instruments are valid instruments and are uncorrelated with the error term can be accepted on 90% confidence interval.

For the firms located in the cities within agglomerations, which are not the agglomeration centers, 'core' and 'llocnl' significant and have inverted U shape; 'lurbnl' is positive and shows that 1% increase in revenue of national firms in the city in different industries increases revenue of national and foreign firms by 0.265%; the coefficient is significant on 95% confidence interval. The coefficients of foreign firms' agglomeration turned out to be insignificant.

The effects for the enterprises located in the monotowns are studied in Model (4). At the 1st stage, for llocf, according to F test of excluded instruments the null hypothesis that the regressor is unidentified can be rejected, but according to Angrist-Pischke multivariate F test of excluded instruments, this hypothesis cannot be rejected. For the variable lurbf, according to both these tests the hypothesis of unidentification can be rejected. At the 2nd stage, Kleibergen-Paap underidentification test rejects the null hypothesis that the model is underidentified. Kleibergen-Paap Wald rk F statistic of 4.134 for weak identification test does not exceed the lowest Stock-Yogo weak ID test critical value of 4.73 for 30% maximal IV relative bias. Hansen J statistic 4.458 with p-value of 0.1076 shows that the joint null

hypothesis that the instruments are valid instruments and are uncorrelated with the error term can be accepted on 90% confidence interval.

'Core' variable has positive and significant effect; the coefficient of 5.443 meaning that 1% increase in the share of revenue in a city in the industry where a firm belongs, increases its revenue by 0.054%; in other words, if this share – 'core' – increases by 0.01, then the revenue of the enterprise will increase in exp(5.433*0.01)=1.0559 times. Coefficient 'llocnl' is positive and significant on 95% confidence interval; if localization of national firms increases by 1%, firm's revenue increases by 0.0183%. While for the other cities and towns the effects of 'core' and 'llocnl' have an inverted U shape, for monotowns they are positive. Concerning urbanization of national firms, if it increases by 1%, firm's revenue increases by 0.224%, meaning that firms working in the other industries have a relatively large impact on the firm's revenue; the coefficient is significant on 95% confidence interval. Concerning HMP arising from the other cities, 1% increase in it leads to the increase in firm's revenue by 0.621%. Policy measures aimed at communications and transportation between the monotowns and other cities, would probably lead to the increase in the firms' revenue there.

The effects for the enterprises located in all other cities and towns are studied in Model (5). F test of excluded instruments and Angrist-Pischke multivariate F test of excluded instruments show that for llocf and lurbf, the null hypothesis that the endogenous regressor is unidentified is accepted. Kleibergen-Paap underidentification test accepts the null hypothesis that the model is underidentified. Kleibergen-Paap Wald rk F statistic of 1.09 for weak identification test does not exceed the lowest Stock-Yogo weak ID test critical value of 4.73 for 30% maximal IV relative bias. Therefore, coefficients are biased, and the results should be treated cautiously. However, Hansen J statistic 0.9 with p-value of 0.6378 shows that the joint null hypothesis that the instruments are valid and uncorrelated with the error term, can be accepted on 90% confidence interval. The effects of share of firms in the same industry ('core') and of localization of national firms have an inverted U shape. The effect generated by urbanization of national firms is positive, while both localization and urbanization effects arising from the foreign firms' concentration in a city are insignificant.

Table 4.17. Regression results for different types of cities; national firms
Dependent variable: D. ln(*revenue*)

	(1) All national firms	(2) Agglomeration centers	(3) Cities within agglomerations, not centers	(4) Monotowns	(5) Other cities
D.llocf	-0.731***	0.900***	-5.125	-0.427	-0.517
	(-2.62)	(2.86)	(-1.04)	(-0.78)	(-1.24)

	(1) All national firms	(2) Agglomeration centers	(3) Cities within agglomerations, not centers	(4) Monotowns	(5) Other cities
D.lurbf	0.613***	-0.608	-18.76	0.705	1.738
	(2.73)	(-1.32)	(-1.20)	(1.14)	(0.68)
D.lfixed_assets	0.182***	0.194***	0.120	0.229***	0.150***
	(14.25)	(7.97)	(0.96)	(6.41)	(7.59)
D.llabour	0.268***	0.215***	0.114	0.188***	0.325***
	(19.87)	(9.11)	(1.17)	(6.23)	(13.68)
D.lhmecity	0.254*	-0.409	-3.919	0.384	0.104
	(1.90)	(-1.29)	(-0.85)	(1.24)	(0.40)
D.lwagecity	0.0684	-1.240*	-0.167	0.142	0.103
	(1.37)	(-1.80)	(-1.05)	(0.78)	(0.78)
D.core	13.83***	-13.14*	13.16**	4.579***	15.61***
	(8.06)	(-1.72)	(2.32)	(3.03)	(6.65)
D.core2	-10.75***	22.05*	-13.29	0.165	-13.23***
	(-6.04)	(1.89)	(-1.14)	(0.17)	(-5.51)
D.llocnl	0.0205***	0.0235	0.0497	0.0201**	0.0179***
	(3.26)	(0.83)	(0.69)	(1.99)	(2.88)
D.llocnl2	-0.00146***	-0.00110	-0.00364	-0.000781	-0.00125***
	(-3.31)	(-0.56)	(-0.88)	(-1.02)	(-2.96)
D.lurbnl	0.267***	-0.250	0.788*	0.220**	0.314***
	(7.31)	(-1.33)	(1.65)	(2.53)	(4.17)
D.a_road	0.000300	-0.000707	-0.0105	-0.000132	0.000398
	(1.03)	(-0.80)	(-1.11)	(-0.26)	(0.65)
D.rw_road	-0.00465*	0.0312**	-0.327	0.00201	-0.00190
	(-1.91)	(2.01)	(-1.20)	(0.45)	(-0.56)
D.lbusnenvrisk	-0.00101	0.0295*	0.231	-0.0265	-0.00648
	(-0.13)	(1.79)	(1.19)	(-1.60)	(-0.63)
_cons	-0.102***	0.423**	1.187	-0.130*	-0.0355
	(-3.57)	(2.48)	(1.04)	(-1.91)	(-0.25)
time fixed effects	yes	yes	yes	yes	yes
N	37830	17306	1567	2092	17218
adj. R^2	-1.340	-3.681	-30.016	0.266	-1.839
sigma_u					
sigma_e	0.842	1.176	2.860	0.406	0.955
rho					
Kleibergen-Paap underidentification test (p-value)	27.940 (0.0000)	35.537 (0.0000)	2.971 (0.3962)	7.443 (0.0590)	2.557 (0.4650)
Kleibergen-Paap					

	(1) All national firms	(2) Agglomeration centers	(3) Cities within agglomerations, not centers	(4) Monotowns	(5) Other cities
weak identification test	5.247	4.460	0.729	1.535	0.473
Hansen overidentification test (p-value)	5.042 (0.0804)	3.786 (0.1506)	0.470 (0.7907)	2.370 (0.3058)	0.228 (0.8923)

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Model (1) was considered in Table 4.15 above. The effects for the enterprises located in the agglomeration centers are reflected in Model (2). At the 1st stage, F test of excluded instruments and Angrist-Pischke multivariate F test of excluded instruments show that both llocf and lurbf are identified, as for each of them the null hypothesis that the endogenous regressor is unidentified is strongly rejected. At the 2nd stage, Kleibergen-Paap underidentification test rejects the null hypothesis that the model is underidentified. Kleibergen-Paap Wald rk F statistic of 4.46 for weak identification test does not exceed the lowest Stock-Yogo weak ID test critical value of 4.73 for 30% maximal IV relative bias. Hansen J statistic 3.786 with p-value of 0.1506 shows that the joint null hypothesis that the instruments are valid and uncorrelated with the error term, can be accepted on 90% confidence interval.

The effects for the enterprises located in the cities within agglomerations but outside agglomeration centers are studied in Model (3). At the 1st stage, for llocf, according to F test of excluded instruments and Angrist-Pischke multivariate F test the null hypothesis that the regressor is unidentified cannot be rejected. For lurbf, according to both these tests the hypothesis of unidentification can be rejected. At the 2nd stage, Kleibergen-Paap underidentification test accepts the null hypothesis that the model is underidentified. Kleibergen-Paap Wald rk F statistic for weak identification test does not exceed the lowest Stock-Yogo weak ID test critical value of 4.73 for 30% maximal IV relative bias. Therefore the results are biased. Hansen J statistic shows that the joint null hypothesis that the instruments are valid and uncorrelated with the error term, can be accepted on 90% confidence interval.

The effects for the enterprises located in the monotowns are studied in Model (4). At the 1st stage, F test of excluded instruments and Angrist-Pischke multivariate F test of excluded instruments show that both llocf and lurbf are identified, as for each of these regressors the null hypothesis that the endogenous regressor is unidentified is strongly rejected. At the 2nd stage, Kleibergen-Paap underidentification test rejects the null hypothesis

that the model is underidentified. Kleibergen-Paap Wald rk F statistic of 1.535 for weak identification test does not exceed the lowest Stock-Yogo weak ID test critical value of 4.73 for 30% maximal IV relative bias. Hansen J statistic shows that the joint null hypothesis that the instruments are valid and uncorrelated with the error term, can be accepted on 90% confidence interval.

The effects for the enterprises located in all other cities and towns are studied in Model (5). At the 1st stage, according to F test of excluded instruments, the null hypothesis that the endogenour regressor llocf is underidentified can be accepted, while with Angrist-Pischke multivariate F test it can be rejected on 95% confidence interval. According to both these tests, the hypothesis that the variable lurbf is underidentified, is accepted. At the 2nd stage, tests reveal underidentification and weak identification. Therefore the results are biased. Hansen J statistic shows that the joint null hypothesis that the instruments are valid and uncorrelated with the error term, can be accepted on 90% confidence interval.

Overall, the results have a pattern similar to that for all tradables. In Models (3) and (5) they should be treated especially cautiously both for all firms producing tradable goods (Table 4.16) and for national firms (Table 4.17). For the firms located in agglomeration centers (Model 2), wage effect turns out to be negative and significant, showing that 1% decrease in wages leads to increase in revenue by 1.240%. It might mean that human capital does not have an effect large enough to overweigh the effect from cheaper labour. 'Core' effect has an inverted U shape in case of agglomeration centers. Localization effect generated by foreign firms remains positive and significant on 99% confidence interval, possibly meaning that national firms located in agglomeration centers are competitive enough compared to foreign firms, and advanced enough to absorb the foreign firms' localization spillovers. The coefficient llocf shows that 1% increase in foreign firms' localization leads to the increase in the national firms' revenue by 0.9%, which is close to the effect for all tradables (0.899%).

The results for the other cities located within agglomerations are doubtful according to the results of the test, probably due to lack of observations (Model 3). Coefficient 'core' is positive (while it has an inverted U shape in case when both national and foreign firms are considered); there is positive effect from urbanization of local firms. For monotowns (Model 4), 'core', 'llocnl' and 'lurbnl' are positive and significant, with the values similar to those for all firms producing tradable goods and on the same confidence intervals. For all other cities and towns (Model 5), the results are similar as well to those for all firms producing tradable goods; these results are doubtful as they are biased according to the tests. The situation in

different types of cities was not analyzed for the foreign firms separately, due to lack of observations. In Table 4.18 below the effects for different groups of industries are considered.

Table 4.18. Regression results for different classes of industries; all tradables
Dependent variable: D. ln(*revenue*)

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
D.llocf	-0.750***	0.252	-0.498	0.341	0.269	-0.340
	(-2.80)	(0.68)	(-1.04)	(0.58)	(0.60)	(-1.32)
D.lurbf	1.362**	0.260	-0.0437	0.658	-1.179	0.280*
	(2.34)	(0.37)	(-0.04)	(0.59)	(-0.65)	(1.83)
D.lfixed_ assets	0.185***	0.195***	0.203***	0.101*	0.202*	0.187***
	(14.17)	(5.86)	(6.62)	(1.80)	(1.87)	(10.32)
D.llabour	0.249***	0.285***	0.217***	0.281***	0.314***	0.266***
	(15.82)	(7.72)	(8.27)	(4.48)	(5.72)	(16.05)
D.lhme city	0.390*	0.593	0.193	0.157	-2.484	-0.0514
	(1.77)	(0.97)	(0.58)	(0.42)	(-1.05)	(-0.22)
D.lwage city	-0.126	-0.0353	0.0882	0.0506	0.154	0.0145
	(-0.99)	(-0.24)	(0.37)	(0.89)	(0.28)	(0.19)
D.core	15.12***	21.76***	12.75***	13.81***	6.182	10.64***
	(7.79)	(6.41)	(6.46)	(5.20)	(1.19)	(4.33)
D.core2	-11.07***	-24.07***	-8.862***	-14.63***	-1.612	-11.30***
	(-6.82)	(-3.20)	(-4.85)	(-3.98)	(-0.34)	(-3.57)
D.llocnl	0.0321***	0.0330	0.0220*	0.0172*	0.0405	0.0487***
	(3.49)	(0.88)	(1.84)	(1.69)	(0.91)	(2.73)
D.llocnl2	-0.00242***	-0.00221	-0.00155	-0.00193***	-0.00434**	-0.00347***
	(-3.40)	(-0.83)	(-1.54)	(-3.03)	(-2.06)	(-2.85)
D.lurbnl	0.320***	0.257**	0.298***	0.392**	0.0979	0.251***
	(5.51)	(2.16)	(3.22)	(2.46)	(0.59)	(4.49)
D.a_road	-0.000892	-0.000427	0.000534	0.000122	-0.000363	0.0000833
	(-1.20)	(-0.29)	(0.32)	(0.21)	(-0.33)	(0.34)
D.rw_ road	-0.00877***	-0.000677	-0.00219	-0.00000556	-0.00529	-0.00734**
	(-2.60)	(-0.08)	(-0.51)	(-0.00)	(-0.62)	(-2.51)
D.lbusn envrisk	-0.0278*	-0.00320	0.00527	-0.0174	0.0208	0.00124
	(-1.76)	(-0.11)	(0.19)	(-0.39)	(0.49)	(0.17)

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
_cons	-0.0818**	-0.0868	-0.120**	0.0528	0.349	0.00160
	(-2.29)	(-1.08)	(-1.98)	(0.23)	(0.88)	(0.04)
time fixed effects	yes	yes	yes	yes	yes	yes
N	41349	4611	10785	5758	658	13510
adj. R^2	-2.775	-0.380	-0.833	-1.251	-0.118	0.052
sigma_u						
sigma_e	1.082	0.611	0.713	0.796	0.695	0.572
rho						
Kleibergen-Paap underidentification test	14.447	1.357	7.696	4.143	2.865	48.968
	(0.0024)	(0.7157)	(0.0527)	(0.2465)	(0.4130)	(0.0000)
(p-value)	2.005	0.342	1.991	1.055	0.656	11.413
Kleibergen-Paap weak identification test						
Hansen overidentification test	0.002	1.530	4.021	8.071	0.738	12.057
	(0.9990)	(0.4654)	(0.1339)	(0.0177)	(0.6915)	(0.0024)
(p-value)						

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Model (1) was considered in Table 4.15 above. For Model (2), at the 1st stage, F test of excluded instruments and Angrist-Pischke multivariate F test of excluded instruments show that for both llocf and lurbf, the null hypothesis that the endogenous regressor is unidentified is accepted. At the 2nd stage, Kleibergen-Paap underidentification test confirms the null hypothesis that the model is underidentified. Kleibergen-Paap Wald rk F statistic for weak identification test does not exceed the lowest Stock-Yogo weak ID test critical value of 4.73 for 30% maximal IV relative bias. Therefore the results are biased. Hansen J statistic shows that the joint null hypothesis that the instruments are valid and uncorrelated with the error term can be accepted on 90% confidence interval. In Model (3), at the 1st stage, F test of excluded instruments and Angrist-Pischke multivariate F test of excluded instruments show that for llocf the null hypothesis of unidentification is rejected. For lurbf this hypothesis can be rejected only according to F test of excluded instruments. At the 2nd stage of estimation, tests show that the model is identified, though the instruments are weak, and bias can occur. Hansen J statistic shows that the joint null hypothesis that the instruments are valid and uncorrelated with the error term can be accepted on 90% confidence interval.

In Model (4), at the 1st stage, only the variable lurbf is identified, only according to the F test of excluded instruments. At the 2nd stage, tests reveal low quality of the model. Therefore the results are biased. In Model (5), at the 1st stage, tests show that variable lurbf is identified, while llocf is not. At the 2nd stage, tests show low quality of the model. Therefore the results are biased. For Model (6), at the 1st stage, both for llocf and lurbf, the null hypothesis that the endogenous regressor is unidentified is strongly rejected. At the 2nd stage, Kleibergen-Paap underidentification test confirms the null hypothesis that the model is underidentified. Kleibergen-Paap Wald rk F statistic for weak identification test exceeds the Stock-Yogo weak ID test critical value of 11.04 for 5% maximal IV relative bias. However, Hansen J statistic shows that the joint null hypothesis that the instruments are valid and uncorrelated with the error term is rejected on 99% confidence interval.

Overall, Models (3) and (6) are of relatively better quality, and therefore are more reliable than the others. Model (3) shows that for basic materials industries, 'core' effect is significant on 99% confidence interval and has an inverted U shape, with lower value than for all tradables. Localization of national firms has positive impact on revenue, and is significant on 90% confidence interval, the coefficient showing that 1% increase in national firms' revenue in the same city and same industry leads to the increase in the enterprise revenue by 0.022%. When all tradable industries are considered (Model 1), this effect has an inverted U shape. Urbanization of national firms has a positive effect, slightly less than for all tradables, with 1% increase in urbanization leading to 0.298% increase in the firm's revenue.

For the science based industries (Model (6)), the effect generated by concentration of foreign firms in different industries in a city is positive and significant on 90% confidence interval, showing that 1% increase in lurbf leads to 0.28% in the firm's revenue. 'Core' effect is significant and has an inverted U shape, being slightly lower than for basic materials industries. Localization of national firms is significant on 99% confidence interval, has an inverted U shape, and is higher than on average for tradable industries (in Model (1)). Urbanization of national firms is positive and significant on 99% confidence interval, with value 0.251, which is slightly lower than on average for tradable firms and for the firms producing basic materials. The other models are not considered as the results are relatively less reliable; however, the effects, in case they are significant, repeat the pattern observed for all tradables. The impact of HMP arising from other cities, and of the regional business environment risk are observed for all tradables (Model (1)), but not for separate groups of industries. The negative effect of foreign firms' localization observed in Model (1) is also lost for groups of industries. In the Table 4.19 the sample consists of national firms.

Table 4.19. Regression results for different classes of industries; national firms
Dependent variable: D. ln(*revenue*)

	(1) All national firms	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
D.llocf	-0.731***	0.277	-0.324	-2.205	0.266	0.248
	(-2.62)	(0.93)	(-0.64)	(-1.23)	(0.59)	(0.77)
D.lurbf	0.613***	0.502	0.0305	-1.114	-1.154	-0.0133
	(2.73)	(1.38)	(0.06)	(-0.63)	(-0.65)	(-0.09)
D.lfixed_ assets	0.182***	0.189***	0.220***	0.175***	0.190*	0.200***
	(14.25)	(5.27)	(5.71)	(2.79)	(1.78)	(10.14)
D.llabour	0.268***	0.312***	0.243***	0.355***	0.266***	0.293***
	(19.87)	(6.81)	(8.86)	(4.06)	(5.13)	(15.91)
D.lhmecity	0.254*	0.720	0.0863	0.572	-0.494	0.137
	(1.90)	(1.36)	(0.37)	(0.63)	(-0.36)	(0.59)
D.lwagecity	0.0684	-0.0669	0.0106	0.0180	0.359	-0.0375
	(1.37)	(-0.55)	(0.17)	(0.17)	(0.57)	(-0.50)
D.core	13.83***	21.96***	14.27***	14.43***	5.681	5.995**
	(8.06)	(7.94)	(7.36)	(2.60)	(0.96)	(2.38)
D.core2	-10.75***	-25.14***	-10.85***	-15.78**	-0.823	-5.513*
	(-6.04)	(-4.25)	(-4.91)	(-2.28)	(-0.16)	(-1.76)
D.llocnl	0.0205***	0.0241	0.0242	-0.0344	0.0303	0.0205
	(3.26)	(1.28)	(1.64)	(-0.72)	(0.58)	(1.24)
D.llocnl2	-0.00146***	-0.00173	-0.00195	0.00209	-0.00348	-0.00145
	(-3.31)	(-1.43)	(-1.57)	(0.60)	(-1.46)	(-1.29)
D.lurbnl	0.267***	0.263***	0.377***	0.396	0.145	0.243***
	(7.31)	(2.72)	(3.86)	(1.54)	(1.37)	(4.40)
D.a_road	0.000300	-0.000674	0.000707	-0.000663	-0.000641	0.0000569
	(1.03)	(-0.57)	(0.96)	(-0.50)	(-0.64)	(0.22)
D.rw_road	-0.00465*	0.000163	0.00113	0.0261	-0.00377	-0.00104
	(-1.91)	(0.02)	(0.32)	(1.38)	(-0.45)	(-0.32)
D.lbusnenv risk	-0.00101	-0.000801	0.00382	0.0434	0.0346	0.01000
	(-0.13)	(-0.04)	(0.17)	(0.65)	(0.67)	(1.45)
_cons	-0.102***	-0.0951	-0.105**	-0.381	-0.0423	0.00921
	(-3.57)	(-1.39)	(-1.97)	(-1.33)	(-0.19)	(0.24)
time fixed effects	yes	yes	yes	yes	yes	yes
N	37830	4187	9629	5289	587	12720
adj. R^2	-1.340	-0.544	-0.055	-15.681	-0.218	0.109
sigma_u						

	(1) All national firms	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
sigma_e	0.842	0.622	0.540	2.135	0.631	0.550
rho						
Kleibergen-Paap underidentification test (p-value)	27.940 (0.0000)	7.281 (0.0635)	10.803 (0.0128)	2.763 (0.4296)	1.922 (0.5887)	41.587 (0.0000)
Kleibergen-Paap weak identification test	5.247	1.947	2.860	0.806	0.425	9.851
Hansen overidentification test (p-value)	5.042 (0.0804)	0.493 (0.7816)	9.073 (0.0107)	5.516 (0.0634)	0.672 (0.7146)	12.514 (0.0019)

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Model (1) was considered in Table 4.15 above. Model (2) is estimated for traditional industries. At the 1st stage, tests show that llocf is identified, according to the Angrist-Pischke multivariate F test of excluded instruments the null hypothesis that the variable is underidentified can be rejected. For lurbf tests show that the null hypothesis that the variable is underidentified can be rejected. At the 2nd stage, tests show that the hypothesis of underidentification can be rejected and that the joint null hypothesis that the instruments are valid and uncorrelated with the error term, is rejected on 90% confidence interval (from Hansen J statistic). However, the instruments are weak (from Kleibergen-Paap rk Wald F statistic), therefore the results are biased.

For Model (3), 1st stage tests show that both for llocf and lurbf, the null hypothesis that the variable is underidentified can be rejected. The 2nd stage tests show that the hypothesis that the equation is underidentified can be rejected (from Kleibergen-Paap rk LM statistic). However, instruments are found to be weak (from Kleibergen-Paap rk Wald F statistic). Besides, Hansen J statistics shows that the joint null hypothesis that the instruments are valid instruments, i.e. uncorrelated with the error term, and that the excluded instruments are correctly excluded from the estimated equation, can be accepted, but only on 99% confidence interval. Therefore, the validity of instruments is doubtful. The results might be biased due to weak instruments, and are not highly reliable due to the possibility of correlation of instruments with the error term.

Model (4) turned out to be underidentified (according to the tests on the 1st and 2nd stage), and instruments were found to be weak. However, Hansen J statistics show that the

joint null hypothesis that the instruments are valid, and that the excluded instruments are correctly excluded from the estimated equation, can be accepted, but only on 95% confidence interval. In Model (5) llocf is underidentified, lurbf is identified; the equation is underidentified, and the instruments are weak. Hansen J statistics shows that the joint null hypothesis that the instruments are valid and uncorrelated with the error term, is rejected on 90% confidence interval. Overall, the results in Models (4) and (5) are not reliable. In Model (6) both llocf and lurbf are identified; the equation is identified too (Kleibergen-Paap underidentification test rejects the null hypothesis that the model is underidentified). Kleibergen-Paap Wald rk F statistic of 9.851 for weak identification test exceeds the Stock-Yogo weak ID test critical value of 7.56 for 10% maximal IV relative bias. However, Hansen J statistics show that the joint null hypothesis that the instruments are valid and uncorrelated with the error term, can be rejected on 99% confidence interval. Therefore, the results are not reliable.

Models (2) and (3) are of relatively good quality. For the national firms working in traditional industries, 'core' effect was found to be much higher than on average for the national firms. It is significant on 99% confidence interval and has an inverted U shape. Urbanization effect arising from presence of the other national firms in the city (variable 'lurbnl'), is significant on 99% confidence interval, and is similar to the 'lurbnl' effect for all national firms. For basic materials (Model (3)), 'core' effect is significant on 99% confidence interval, is slightly higher than for all national firms on average, and has an inverted U shape. Urbanization effect arising from presence of other national firms in the city, is significant on 99% confidence interval, and is higher than the 'lurbnl' effect for all national firms. The other models are not analyzed as the results are relatively less reliable; however, the effects, when they are significant, repeat the pattern observed for all national firms. In Table 4.20 below the effects for national firms of different age, productivity and ownership are analyzed.

Table 4.20. Regression results for national firms of different age, productivity and ownership

Dependent variable: D. ln(*revenue*)

	(1) Young	(2) Old	(3) More productive	(4) Less productive	(5) Private	(6) State
D.llocf	-1.255***	-0.532*	0.0500	-0.952***	-0.552*	-0.847
	(-2.61)	(-1.75)	(0.13)	(-3.16)	(-1.89)	(-1.42)
D.lurbf	1.499**	0.413*	0.107	0.674**	0.559**	0.559
	(2.42)	(1.86)	(0.43)	(2.53)	(2.15)	(1.28)
D.lfixed_ assets	0.177***	0.181***	0.136***	0.184***	0.185***	0.209***
	(8.14)	(10.11)	(6.84)	(12.09)	(12.59)	(8.45)

	(1) Young	(2) Old	(3) More productive	(4) Less productive	(5) Private	(6) State
D.llabour	0.276***	0.269***	0.172***	0.310***	0.257***	0.266***
	(12.33)	(15.98)	(8.46)	(18.65)	(17.74)	(9.36)
D.lhmecity	0.451	0.226	-0.141	0.236	0.188	0.652
	(1.45)	(1.37)	(-0.51)	(1.54)	(1.54)	(1.34)
D.lwagecity	0.00766	0.0847	-0.0200	0.156**	0.0293	0.131
	(0.11)	(1.10)	(-0.65)	(2.02)	(0.67)	(1.04)
D.core	15.86***	13.23***	8.661***	17.31***	12.78***	16.57***
	(7.63)	(6.01)	(7.85)	(7.46)	(7.81)	(2.89)
D.core2	-11.46***	-11.40***	-4.866***	-16.85***	-9.505***	-14.52**
	(-5.52)	(-4.72)	(-4.44)	(-6.17)	(-5.82)	(-2.24)
D.llocnl	0.0347***	0.0187**	0.0204***	0.0289***	0.0210***	0.0345
	(2.82)	(2.16)	(2.63)	(3.43)	(3.79)	(1.46)
D.llocnl2	-0.00241***	-0.00124**	-0.00155**	-0.00216***	-0.00139***	-0.00272
	(-3.46)	(-2.03)	(-2.28)	(-3.53)	(-3.57)	(-1.51)
D.lurbnl	0.351***	0.238***	0.327***	0.236***	0.289***	0.282***
	(5.14)	(5.13)	(4.38)	(5.73)	(7.07)	(2.76)
D.a_road	0.000686	0.000191	-0.000260	0.000439	0.0000328	0.000456
	(0.92)	(0.62)	(-0.91)	(1.18)	(0.12)	(0.65)
D.rw_road	-0.00872	-0.00230	-0.00547	-0.00574**	-0.00177	-0.00933
	(-1.60)	(-0.84)	(-1.18)	(-2.08)	(-0.78)	(-1.54)
D.lbusnenvrisk	-0.0107	0.0000981	-0.00510	-0.00167	-0.00791	0.0233
	(-0.61)	(0.01)	(-0.53)	(-0.17)	(-0.88)	(1.14)
_cons	-0.122**	-0.0811**	0.140***	-0.169***	-0.0878***	-0.163
	(-2.07)	(-2.37)	(2.84)	(-4.81)	(-3.39)	(-1.58)
time fixed effects	yes	yes	yes	yes	yes	yes
N	12345	25485	6284	31546	28170	14574
adj. R^2	-2.962	-0.865	0.195	-2.205	-0.798	-1.716
sigma_u						
sigma_e	1.238	0.698	0.442	0.997	0.728	0.947
rho						
Kleibergen-Paap underidentification test (p-value)	10.639 (0.0138)	16.870 (0.0008)	10.065 (0.0180)	21.268 (0.0001)	21.806 (0.0001)	15.497 (0.0014)

	(1) Young	(2) Old	(3) More productive	(4) Less productive	(5) Private	(6) State
Kleibergen-Paap weak identification test	2.438	3.322	2.437	3.741	3.325	3.789
Hansen overidentification test (p-value)	1.481 (0.4770)	5.112 (0.0776)	0.418 (0.8112)	4.419 (0.1097)	11.635 (0.0030)	6.301 (0.0428)

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In Models (1), (2), (3), (4) and (6) llocf and lurbf are identified; the equation is identified, instruments are weak, but uncorrelated with the error term. In Model (5) llocf and lurbf are identified; the equation is identified, instruments are weak, and the hypothesis that they are uncorrelated with the error term is rejected. In other words, Model (5) is relatively less reliable. In the other models coefficients are biased due to weak instruments, though the value of Kleibergen-Paap rk Wald F statistic is relatively slightly lower than the critical value for 30% maximal IV relative bias (4.73).

Localization effect arising from foreign firms is either negative or insignificant in the models considered above. It is significant on 99% confidence interval for the young firms, and shows that 1% increase in foreign firms' localization leads to 1.255% decrease in revenue for the 'young' national firms. For the 'old' national firms this coefficient is significant on 90% confidence interval and shows that 1% increase in foreign firms' localization leads to 0.532% decrease in 'old' national firms' revenue. For the firms with above average productivity the coefficient is insignificant, possibly meaning that they are able to compete with foreign firms in the same industry; however they are not found to benefit from presence of foreign firms in the same industry. For the firms with below average productivity 1% increase in revenue of foreign firms in the same industry leads to decrease in their own revenue by 0.952%; the coefficient is significant on 99% confidence interval. As for private firms, their revenue decreases by 0.552% with 1% increase in localization of foreign firms; the coefficient is significant on 90% confidence interval. For state firms this coefficient is not significant.

Concerning urbanization of foreign firms, it is either positive or insignificant for various types of firms. For the 'young' firms the coefficient is the highest, showing that these firms' revenue increases by 1.499% with 1% increase in urbanization of foreign firms. For 'old' firms the coefficient is lower, showing that 1% increase in foreign firms' urbanization

leads to 0.413% increase in revenue of 'old' national firms. It might reflect the concept of the product life cycle, when urbanization economies are relatively more important for younger firms, though in this case the division between the 'young' and 'old' firms is relative. Some 'young' firms might have been established on the premises of the 'old' firms during privatization. Besides, below it will be discussed that 'young' firms were found to benefit relatively more from localization economies (arising from national firms). As for productivity levels, foreign firms' urbanization turned out to be significant for relatively less productive firms, showing that 1% increase in urbanization leads to 0.674% increase in the firm's revenue. Private firms were found to benefit from foreign firms' urbanization, with 1% increase in urbanization leading to 0.559% increase in their revenue.

HMP arising from the other cities is insignificant, when types of firms are considered separately, although it is significant when the model is estimated for all national firms. Wage effect turned out to be positive and significant for less productive firms. It seems counter intuitive, implying that for less productive firms human capital is relatively more important. The effects generated by the share of firms in the same industry in the revenue of all firms, reflected in 'core' variable are significant and have inverted U shape for all types of firms, being relatively higher for 'young' firms, less productive firms and state firms. The effects arising from localization of national firms are significant for all types of firms except for state firms and have inverted U shape, being relatively higher and more significant for 'young' firms, and slightly higher for less productive firms than for more productive ones. It implies that younger firms and private firms are relatively more susceptible to agglomeration economies arising from concentration of firms in the same industry. Urbanization effects generated by national firms were found to be positive and significant on 99% confidence interval for all types of firms, and are relatively larger for 'young' firms, more productive firms, and slightly larger for private than for state firms.

The effects arising from concentration of firms with foreign ownership are likely to be biased downward for two reasons. Firstly, the share of foreigners in the ownership of the companies is not known; firms with any share owned by foreigners are classified as 'foreign'; in certain firms the share might be low. Secondly, the origin of investment registered as foreign is not known. There might be some 'foreign' firms actually belonging to the Russian owners registered in the foreign countries. Besides, a more precise analysis is needed to reflect vertical, horizontal and supply-backward FDI effects. Vast literature exists on this matter, a brief overview of it is provided in Chapter 1.

4.6. Conclusion

The objective in this chapter was to reveal the impact of business activity concentration on enterprise productivity, paying special attention to the effects arising from concentration of national and foreign business. Descriptive statistics on agglomeration indices shows that foreign firms tend to locate in places with higher diversity levels, arising both from national and foreign firms. They are attracted to the cities, where the other foreign firms are already located, presence of foreign firms working in the other industries having high importance, and presence of national and foreign firms working in their industry being important as well.

The first question, posed in this chapter was: 'what are the effects on enterprise performance generated by concentration of national and foreign enterprises?'. The first hypothesis that both national firms' urbanization and HMP positively affect the enterprise productivity was confirmed. It was also confirmed that this impact varies across industries and cities; for some subsamples it was found to be insignificant. The second hypothesis that localization effects have an inverted U shape was confirmed too. The third hypothesis that FDI concentration may have controversial localization effects, but positive urbanization effects, was confirmed.

The following models were considered: the model with fixed effects and without instruments, the model with fixed effects and with instruments, the model with first differencing and with instruments. The model with fixed effects and without instruments led to the following results. As for the agglomeration externalities generated both by national and foreign firms, localization externalities are significant and have an inverted U shape. Urbanization externalities are significant and positive. Results show that firms with foreign ownership benefit more both from localization and urbanization externalities. The reasons behind it might be that they have more choice where to locate, and that their level of human capital, particularly of management, might be higher. However, this result is not robust to subdividing agglomeration economies into those generated by national and foreign firms; only localization effects generated by foreign firms are much stronger for the foreign firms than for all firms producing tradable goods; they are positive.

Overall, in the model with fixed effects it was found that localization externalities generated by foreign firms are significant and positive for the other foreign firms, i.e. firms with foreign ownership benefit from presence of foreign firms in the same industry. Localization effects generated by national firms are significant and positive for the other national firms (without an inverted U shape). The reason behind it might be that national and

foreign firms working in the same industry and the same city have to a certain extent different labour markets, different suppliers and do not interact intensively enough to share knowledge.

In the model with firm fixed effects and GMM, for all tradables, localization effects arising from national firms were found to have inverted U shape, localization effects arising from foreign firms are insignificant. Urbanization effects arising both from national and foreign firms are significant and positive, foreign firms' urbanization effects being much higher than national firms' ones. The foreign firms' urbanization level is important probably because existence of foreign firms is favourable for development of suppliers for enterprises belonging to various industries working in a city, such as business services. Besides, by bringing in new business practices they contribute to formation of favorable business climate. The effect of HMP arising from other cities stayed positive and significant. Business environment risks produce negative effects.

In the model with first differencing and GMM, the effect of urbanization arising from national firms is robust to the change in the estimation technique. Urbanization effects arising from foreign firms are significant and have high value. Localization effects arising from national firms, keep the inverted U shape; this effect is lower than in the model with firm fixed effects and GMM. Localization arising from the foreign firms produces a negative effect on firms' productivity in this model, while in the model with firm fixed effects it was insignificant. HMP remains significant and positive for all firms and for national firms, with lower value than in the model with firm fixed effects and GMM. The results are similar in the models with firms fixed effects and GMM, and with first differencing and GMM.

The next two hypotheses are the same as in the previous chapter. The fourth hypothesis states, that regional transport infrastructure and business climate are important for enterprise performance. The fifth hypothesis is that human capital level is also important for enterprise performance.

The model with fixed effects showed that regional business environment risks negatively affect firm performance; this result is robust for the majority of specifications, and is especially important for the foreign firms (Model 5 in Table 4.10). In the model with first differencing and GMM, business environment risks stay significant in the specification with all tradables, and keep negative sign. The logic behind this indicator is that business owners and managers have to spend resources to deal with the difficulties of economic environment rather than to improve their products, production processes and management practices.

The effect of automobile road density was significant and positive in all specifications with fixed effects and without instruments (Models 1 – 6, Table 4.10), while the effect of railroad density was either negative or insignificant, possibly because the railroad

infrastructure is not developed enough. The effects of infrastructure were not revealed clearly in the models with fixed effects – with and without instruments, and in the model with first differencing and instruments. Wage effect turned out to be negative and significant in the model with fixed effects and instruments, and insignificant in the model with first differencing and instruments. Therefore, the effect of human capital was not revealed, probably due to the imperfect proxy, city average wage, used for this purpose.

Another question was 'How do agglomeration effects depend on the enterprise characteristics, industry, and city type?' Concerning enterprise characteristics, according to the models with fixed effects, the national firms seem to benefit more from interaction with the other national firms. However, national firms do benefit from externalities generated by foreign enterprises as well. Localization externalities generated by foreign firms proved to be beneficial for the national firms in science based industries. Urbanization economies generated by foreign firms proved to be positive and significant for the national firms belonging to science based and traditional industries. Localization externalities generated by foreign firms are significant for the other foreign firms belonging to science based industries and have an inverted U shape; urbanization externalities generated by foreign firms are significant for the other foreign firms belonging to basic machinery. Urbanization externalities generated by national firms are significant for the foreign firms belonging to basic machinery and science based industries. Overall, urbanization economies both arising from foreign firms and from national firms are beneficial both for national and foreign firms. However, localization economies hold mostly for the firms of the same type.

For the models with first differencing and GMM the results are the following. Comparing national firms with the sample for all tradables, the results have similar pattern, i.e. 'core' and 'llocnl' have inverted U shape; 'lurbnl' and 'lurbf' are positive, while 'llocf' are negative when all firms are considered. HMP of the other cities has positive effect. Business environment risks have negative effect when all tradables are considered, and are insignificant for the national firms. Agglomeration effects arising from foreign firms' presence in a city are relatively more important for 'young' firms and 'private' firms. Foreign firms' presence in different industries in a city (urbanization) is beneficial for the national firms of different types, while competition effects from their presence in the same industry (localization) outweighs positive agglomeration effects. Share of firms in the same industry ('core') is relatively more important for 'young', state, and less productive firms. Localization of national firms is more important for 'young', private and less productive firms. Urbanization effects generated by national firms are positive and significant for all types of firms, are relatively larger for 'young' and more productive firms, and are slightly larger for

private than for state firms. On one hand, these firms are probably more involved in market mechanisms, and therefore are affected by the other enterprises. On the other hand, they might be able to benefit more from agglomeration effects due to more efficient management.

The effects for the firms located in various types of cities are analyzed in the model with firm fixed effects and GMM. For all firms producing tradable goods located in different city types, the results are the following. Localization effects arising from national firms are significant and keep the inverted U shape for all cities and towns, which are not monotowns and not agglomeration centers. For monotowns national firms' localization effect is positive, while for agglomeration centers it is not significant. The 'core' effect is also positive for monotowns; for agglomeration centers it is insignificant, and for the other cities and towns has an inverted U shape. National firms' urbanization effect is positive and significant for all towns and cities, except for agglomeration centers. Foreign firms' localization effects are positive and significant for agglomeration centers; for the other cities they are insignificant, though when all firms producing tradable goods are considered together, these effects are significant and negative. Foreign firms' urbanization effects become insignificant for all city types, though they are positive and significant for all firms producing tradable goods considered together. Positive effects of HMP arising from the other cities, for monotowns, show that improvement of communication and transportation between the monotowns and the other cities, would lead to increase in revenue of firms there.

As for the national firms located in different city types, the results are similar to those considered above. The most reliable models were estimated for agglomeration centers and for monotowns. For agglomeration centers, localization effects generated by foreign firms are positive and significant, possibly meaning that national firms located in agglomeration centers are competitive enough compared to foreign firms, and advanced enough to absorb foreign firms' localization spillovers. The effect arising from the share of the firms in the revenue of the same industry in the city, 'core', has inverted U shape. For monotowns, national firms' urbanization effects are positive and significant; national firms' localization effects and 'core' are also positive and significant, all coefficients having the values comparable to those for all firms producing tradable goods. For the firms located in agglomeration centers, wage effect turns out to be negative and significant. It might mean that human capital does not have an effect large enough to overweigh the effect from cheaper labour. HMP arising from other cities, loses significance for the sample of national firms.

Concerning groups of industries, when all firms producing tradable goods are considered together, foreign firms' localization effects are insignificant. Foreign firms' urbanization effects are significant and positive for science based industries, showing that 1%

increase in 'lurbf' leads to 0.28% increase in the firm's revenue. 'Core' is significant and retains the inverted U shape for all industries but integrating machinery, where it is insignificant. National firms' localization effects are positive for basic materials, have inverted U shape for basic machinery and science based industries, and are not significant for traditional and integrating machinery industries. National firms' urbanization effects are significant and positive for all industries but integrating machinery.

As for the national firms, the effects from the foreign firms' presence are observed for all national firms, but not for groups of industries considered separately. The impact of HMP arising from the other cities is also observed for all national firms, but not for separate groups of industries. The same is true about localization effects generated by the national firms, possibly due to lack of observations. For groups of industries, the effect from the share of the same industry in the city's revenue ('core') is observed, for all industries but integrating machinery. Urbanization effects generated by national firms' presence are significant and positive for traditional, basic materials, and science based industries. However, for all tradable industries (national and foreign firms together) considered above, effects from foreign firms' urbanization and national firms' localization were observed for separate groups of industries. HMP, wages, road infrastructure and business environment effects are not revealed for separate groups of industries – both for all tradables and for the national firms. However, there is significant positive effect of HMP arising from the other cities when all tradables or all national firms are considered together. There is also significant negative effect of business environment risk when all tradable industries are considered together.

5. Interaction of agglomeration economies and home market potential in the firm's location choice

5.1. Introduction

In the previous chapters the effects of spatial concentration of economic activities on enterprise performance were examined. This chapter is devoted to the determinants of enterprise location choice for cities. There is a theoretical framework explaining circular causality of agglomeration process, when the existing agglomeration externalities attract new firms, and the more firms there are in a territory, the higher the externalities become (Ottaviano, Thisse, 2004). If to assume that the cumulative agglomeration process can be caused by a small change in production factors, then the regional policy attracting capital and labour into a less successful territory would lead to increased interest of firms in this territory.

It would happen due to increasing returns for enterprises even after the policy intervention is ended. However, there is a question if regional policy has a long run effect. The effects are long run under assumption that multiple equilibria exist, i.e. that any of several symmetric territories can become an agglomeration if there is a policy intervention (Mikhailova, 2011).

In the Soviet Union, social necessity and political factors rather than economic efficiency were taken into account while choosing enterprise locations (Mikhailova, 2011). Besides, population density is comparatively low in Russia, while territory is large. It can explain why there are relatively few agglomerations, 'development centers', in a large country as Russia.

Concerning 'agglomeration centers', on one hand, they drive economic growth and facilitate development of the other territories as well. On the other hand, excessive inequalities between territories can lead not only to social instability, but also to human capital decline, undermining economic growth. The dilemma of regional equality vs. economic growth became especially acute during the recent economic crisis (Shiltsin, 2010). If a certain territory develops, the other territories benefit as well, if some resources are reallocated to them. Then the aim of economic policy can be twofold. On one hand, to support development of agglomeration centers with regional policy instruments, and on the other hand, to sustain and improve education, medical care and the other socially significant areas in the lagging territories with social policy instruments (Martin, 1998). Moreover, economic policy measures can also support long term development of the lagging regions, particularly, through promoting investment projects (Shiltsin, 2010).

On the city level, taking into account competition between cities for financial and human resources, success of a city depends on whether firms choose to work there. City attractiveness for enterprises is interrelated with city attractiveness to human capital, as it is reflected in theoretical models such as those with footloose entrepreneurs (Forslid, Ottaviano, 2003). Enterprise location choice has also social meaning, being closely linked to availability of well paid jobs and high standards of living.

Based on enterprise level data for Russia, and city and regional data linked to it, this research studies the determinants of enterprise location choice for cities. The objective is to find out the city characteristics favorable for business location, particularly those that can be affected with economic policy. Specific features of industries, cities and firms are considered. Location choices of national and foreign enterprises are analyzed. For national firms it rather implies the probability of formation and work in a city, while for the foreign firms it actually is location choice among cities.

The following questions are considered in this study. Which characteristics of cities are important for attracting business? Do enterprises tend to locate in cities with higher agglomeration levels and home market potential (HMP), i.e. do agglomeration forces prevail or there is a tendency for more even distribution of economic activity? How do location choice determinants vary across industries? Which characteristics of cities are relatively more important for different types of firms: national and foreign firms, firms with different productivity levels, 'young' and 'old', private and state firms?

Comparative advantage of a city is assumed to be formed by costs, resource availability and agglomeration effects, i.e. firms choose cities where resources are available at lower cost (lower wages, developed infrastructure etc.). It is assumed that among the factors of enterprise choice for a city are agglomeration level, HMP, costs, human capital, transport infrastructure, and business climate (high investment potential and low investment risk). For the urban growth, skills are essential, particularly in the areas of the country with colder climate (Glaeser, Gottlieb, 2009). Yet, the more skilled the workers are in a city, the higher are wages, so costs of firms are higher (Glaeser, Gottlieb, 2009), i.e. there is a trade-off between labor costs and human capital quality.

The first hypothesis is that firms are attracted to the cities with sufficient economic activity of the other firms of various industries, i.e. urbanization and diversity effects are important. The second hypothesis is that firms are attracted to the cities with sufficient economic activity of the other firms of the same industry, i.e. localization effects are present. Localization effects are assumed to have inverted U shape. The third hypothesis is that HMP is significant (based on Krugman's model (1980)). Firms choose cities where they will have a better market access, i.e. cities with large market size of their own and with relatively less costly access to the other cities with large market size. The fourth hypothesis is that wages ambiguously affect location choice (Head, Mayer, 2004).

Infrastructure can be subdivided into physical and institutional aspects (Kolomak, 2006). The fifth hypothesis is that physical infrastructure development measured with automobile and railroad density positively affects city attractiveness. The sixth hypothesis is that firms choose cities located in the regions with lower business environment risks, i.e. developed institutional infrastructure. Besides, it is assumed that location advantages and type of city matter: incentives of firms to locate in agglomeration centers, other cities belonging to the agglomerations, and monotowns differ[46]. It is assumed that foreign firms are more flexible in their location choice, so their choice for a city is more affected by location determinants.

[46] Definition from Vorobyev et al. (2010). More detail is in Chapter 2.

The chapter is organized as follows. In section 5.2, review of empirical results is presented. In section 5.3, theoretical background of agglomeration economies and HMP is reviewed. In section 5.4, empirical background for specification of location choice model is considered. In section 5.5, estimation of the model is discussed. In section 5.6, data is presented. In section 5.7, the results are discussed, and the conclusion follows.

5.2. Review of empirical results

Some facts on location choice of foreign enterprises in Russia. In this chapter location choice for a city made by national and foreign firms is analyzed. Location choice of foreign enterprises is of interest for two reasons. Firstly, they have more opportunities to choose a city than national firms do. Secondly, foreign direct investment (FDI) not only produces spillovers in terms of productivity growth for the other firms, but also has an impact on institutions and managerial skills thus forming development perspectives in a transition country (Castiglione et al., 2012). Concerning FDI in Russia, formation of joint ventures with foreign companies became possible starting from 1988[47]. Since then, Russia's official credibility rating improved; government managed to restore common legal space; in 2000s credit conditions were favourable; economic growth was about 7% per year in 1999-2004 (Ahrend, 2006). All this made Russia attractive to FDI. Russia was the third largest FDI recipient among the countries of Central and Eastern Europe, the first among the former republics of the Soviet Union.

The share of global inward FDI flow in Russia exceeds the share of global exports and imports, the share of FDI being 3.5%, while the share of world merchandise exports is 2.9% and of imports is 1.8%, according to *WTO* (2012) and *World Investment Report* (UNCTAD, 2012) (Gonchar, Marek, 2013). FDI inward stock in Russia in 2000 was $32,204 mn, it reached $423,150 mn in 2010, and $457,474 mn in 2011. For comparison, total FDI inward stock in all countries of the Commonwealth of Independent States (CIS), including Russia, was $55,159 mn in 2000, $611,418 mn in 2010 and $672,253 mn in 2011 (World Investment Report, UNCTAD, 2011 and 2012).

In 2009-2010 Russia ranked the 7th out of top 20 host economies in terms of global FDI inflows, in 2009 FDI inflows being $36 bn and in 2010 – $41 bn (World Investment Report, UNCTAD, 2011). However, the *World Investment Report* (UNCTAD, 2003) attributed to Russia the rank 108 out of 140 countries in the UNCTAD inward FDI

[47]Database of legislation documents 'Garant' http://base.garant.ru/10103075/ (The Law of USSR 'On cooperation in USSR' from 26 May 2008 # 8998-XI. Includes regulation of joint ventures – article 28 'Foreign economic activity of cooperatives', clause 4)

performance index for 1999-2001; this index measures the amount of FDI that countries receive relative to their GDP. FDI inward stock as a percentage of GDP in Russia was 12,4% in 2000, 33% in 2010 and 25,7% in 2012, while in CIS it was 15,6%, 34,5% and 29,1% in the same years, and in the European Union – 27,7%, 45,8% and 46,6% (UNCTAD, 2013).

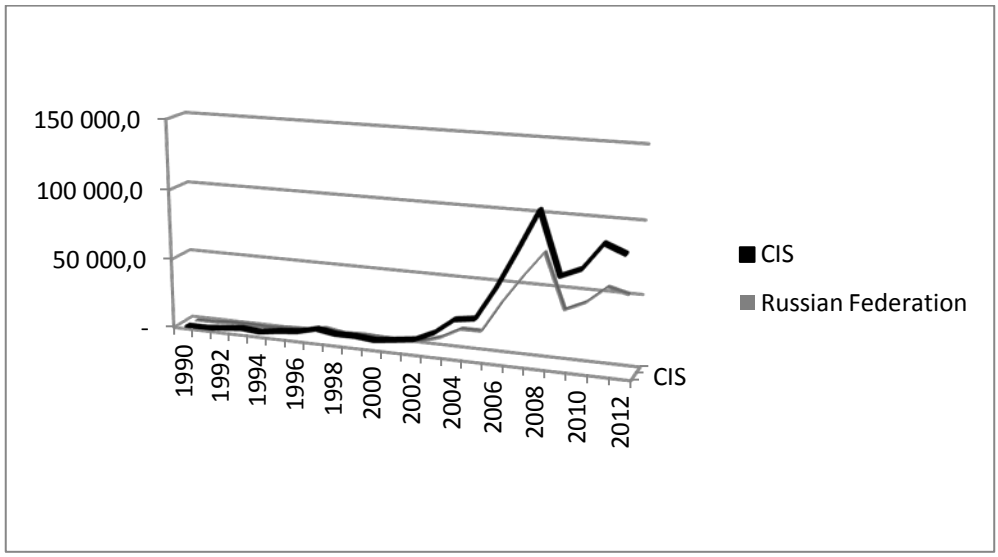


Figure 5.1 FDI inflows, in the Commonwealth of Independent States, including Russian Federation, 1990-2012
Source: World Investment Report 2013: Annex Tables, UNCTAD

By the year 2008, 78 subjects of the Russian Federation (out of 83) had regional laws on state regulation or support of investments, uniting all preferences and limitations for investors. The other regions have laws on separate aspects of investment activities. According to the legislation of the Russia Federation, legal regime for the foreign investments cannot be different from that for the residents. However, several regions have passed separate laws regulating foreign investments during the 1990ies (Bashkortostan, Kalmykia, Tatarstan, Buryatia, Sakha, Stavropol'sky, Permsky, Zabaykal'sky krai, Orenburg region); according to the experts, such laws do not contribute to the attractiveness of the region (Expert, 2008). Concerning FDI in particular industries, for example, a new special economic zone was created in the Samarskaya Region in order to attract investors particularly in the car-making and related industries. Besides, simplified rules were introduced for employing highly qualified foreign specialists[48] (UNCTAD, 2011).

In the beginning of transition process there was lack of consumer goods (as a result of consumer goods deficit existing in the planned economy), accompanied by high inflation. Therefore consumption was preferred to savings, and foreign companies were motivated by the large market potential. Besides, income grew, and there was rouble appreciation in the

[48] Federal Law No. 86-FZ, 19 May 2010.

mid of 1990ies. Another aspect that attracted foreign investors was natural resources endowment (Castiglione et al., 2012).

There are significant disproportions of spatial development in Russia. In 2008 Moscow accounted for 23% of total Gross Regional Product of Russia. 35% of total GRP were produced jointly by Moscow and Tyumen oblast, including its oil-extracting districts ('okrugs'). Ten largest regions of Russia, out of 83, generate 57% of total GRP. GRP per capita in Tyumen oblast is 4 times average GRP per capita in Russia; GRP per capita in Moscow is 2 times average GRP per capita in Russia; ten regions produce above average GRP per capita[49].

FDI is also distributed unevenly across regions, and across sectors (Castiglione et al., 2012). Ledyaeva (2009) points out ten Russian regions - leaders in attracting FDI; in 1999-2006 they were Sakhalin, Moscow, Omsk region, Moscow region, Tyumen region, Lipetsk region, Leningrad region, Krasnodar region, Novgorod region, St. Petersburg. According to Castiglione et al. (2012), in 2001 out of 79 regions 6 regions received 70% of FDI: Moscow city (28.3%), Krasnodar region (16.82%), Sakhalin (9.17%), Moscow region (7.66%), Leningrad region (5.84%), St. Petersburg city (2.8%). Another 12.2% were received by Samara, Novosibirsk, Sverdlovsk, Tyumen, Orenburg regions. 61 regions received the remaining 17% of foreign investment, while 17 regions did not receive any FDI.

FDI is affected by numerous region specific factors: development of infrastructure, business services, such as financial sector, consulting, by rent, prices etc. Overall, the number of firms with FDI varies significantly across regions and cities. In Figure 5.2 below there is spatial distribution of firms with FDI across the Russian regions (Gonchar and Marek, 2013).

[49] Zubarevich (2008) Strategies of long term development. Remember about space. *Vedomosti*, 4 September, p. 4.

Source: Sample data, Greenfield and M&A investments are taken into account

Figure 5.2 Spatial distribution of FDI firms in the sample**:** FDI firms in the region as a percentage of the total number of foreign firms in the country

Source: Gonchar, Marek (2013), based on data for 2000-2009

Bradshaw (2002) offered the following classification of the Russian regions and territories according to the investors' interest. The first group is Moscow city and Moscow region as the centers of political power. The second group is regional, financial and industrial centers associated with large local markets (Krasnodar, Leningrad regions, Novosibirsk, St. Petersburg, Samara, Sverdlovsk). The third group is regions having sea ports or situated near the frontier (Krasnodar, St. Petersburg, Far East). The fourth group is energy producing regions (Sakhalin, Tyumen). The fifth group is regions with import-substitution industries (Moscow, St. Petersburg).

Bradshaw (2002) cites the results of the investors' surveys pointing out the factors making investment in Russia difficult: high tax burden and complicated tax system; lack of international accounting standards; unclear and over bureaucratic system of standards; crime and corruption; inadequate protection of property rights; problems with customs and check points. He also emphasizes lack of political and economic stability and cites the survey of Economist Intelligence Unit conducted in 2001 according to which Russia ranked 5.5 out of 10, while Poland and Czech Republic ranked 7.0, and Hungary ranked 7.2, according to political environment and expected policies towards investments in 2001-2005.

Empirical research. Head and Mayer (2004) combine the aspects of market size and agglomeration externalities. They study HMP using a theoretical model of location choice that gives a profit function containing market potential index introduced by Harris (1954), spatial distribution of competitors and agglomeration variables. Based on Japanese firm level

data for 1984-1995 they find that location choice is affected both by market potential and by agglomeration forces. Combes et al. (2008a) present the model by Krugman (1980, 1991), the overview of its empirical testing, and analysis of the work by Head and Mayer (2004).

Concerning enterprises with FDI, Crozet et al. (2004) conclude that regional policy does not significantly contribute to attraction of FDI. However, the experience of some countries, for example, Ireland showed that economic policy and state institutions affect development of projects with participation of FDI. Ireland chose a strategy of attracting the firms from the sectors with medium economy to scale and low transport costs, with a tendency to reallocation from the European economic centers to periphery regions with highly qualified workforce, such as office equipment, radio- and telecommunications (Navaretti, Venables, 2004). Barry et al. (2001) find that agglomeration and demonstration effects are significant for foreign business in Ireland. Agglomeration effects are more important for high tech industries; both effects are equally important for low tech industries, and both effects are particularly strong for the firms from the same country.

Investment climate is among the factors essential for the enterprise location choice. Shastitko et al. (2007) claims that investment climate is formed by economic, political and social components. Along with investment climate, investment image, i.e. investors' opinions on various aspects of investment climate is important to attract firms. It is reasonable for government and local business to form realistic investment image of a region or country.

Evidence for transition countries. Boudier-Bensebaa (2005) studied the factors determining location choices of enterprises with FDI among the Hungarian counties. The dependent variable was FDI stocks in a county in 1991-2000. She found that labor availability, industrial demand, manufacturing density, inter-industrial agglomeration economies and infrastructure are significant in attracting FDI. The results showed that foreign firms are attracted to the regions with higher labor costs.

There is a number of economic growth studies and enterprise location choice studies on the city level for transition countries. Belitsky (2011) uses dynamic panel data on former Soviet Union cities for 1995-2008. He estimates variation in GDP per capita growth rates in 98 cities with System GMM technique and finds that factors affecting cities economic growth are size of the local government, level of poverty and quality of institutions, i.e. property rights protection and freedom of doing business. Shepotilo (2011) studies the countries of Europe and Central Asia region. According to him, in Soviet Union city development was distorted by central planning, therefore population distribution in the country is more even than it would have been under market economy. He concludes that now population mobility

remains low, and convergence to a new spatial equilibrium is low due to underdeveloped cities infrastructure and inconsistent government policy.

As for the effect of policy intervention on city development, (Mikhailova, 2011) studies the evolution of city growth in the Soviet Union and Russia in the 20th century with the aim to find out if multiple spatial equilibria exist. Transformation from centrally planned system towards the market economy gave an opportunity to follow the adjustment to a market based spatial equilibrium. She separates the impact of location fundamentals and agglomeration externalities on regional growth. Her question is if there is mean reversion in the city growth during the Soviet Union and transition period, and if the spatial process is in line with single or multiple equilibria model. The results are consistent with multiple equilibria hypothesis. Statistically significant consequences of the Second World War (destruction-type shocks) for city growth were not found, controlling for the other variables. It was found that the GULAG system (relocation-type shocks) led to long term consequences, i.e. the cities that received additional production factors within this system in 1930-1950ies, had a higher chance to continue growing, and the cities that did not, were likely to decline.

Concerning location choice of the foreign firms, the common framework of analysis is to classify FDI into investment motivated by market access ('horizontal') and investment motivated by production costs differentials ('vertical'). There are also frameworks integrating these two main motivations (Mayer et al., 2010). 'Horizontal' FDI is investment into production facility abroad with the aim to serve a foreign market, while 'vertical' investment implies building a production chain that involves facilities in several countries (Helpman et al., 2004).

Strategies of multinational firms in decision making about FDI location are classified into the traditional ones (resources seeking, market seeking, asset seeking, efficiency seeking), and more complex global strategies nowadays followed by the majority of firms (Andreff, 2003). Resources seeking strategies are associated with allocation of enterprises in the regions with larger natural resources endowment. Market seeking strategies account for interest of enterprises in large regional demand. Asset seeking strategies reflect interest in technologically advanced production capacities. Efficiency seeking strategies imply allocation in places where costs are lower, for example, regions with lower wages (Dunning, 1993).

To account for market seeking strategies, market size (measured as gross regional product, gross regional product per capita, or population) is included by researchers as an explanatory variable (Castiglione et al., 2012). Taking into consideration large size of Russia, researchers account for geographical factors, such as distance from regional capitals to Moscow and presence of coastal borders.

Concerning efficiency seeking investments, production costs differences are essential. However, as for labour costs, if they are low it means that income per capita is low, so potential demand is low, not mentioning the possibility of lower level of human capital. Association of labour costs and human capital level in Russia is reasonable but disputable. On one hand, there is low labour mobility, so firms can pay lower wages to employees in smaller towns with less opportunity to change a job. Therefore the tendencies for lower wages in smaller towns discussed in economic geography models are even more pronounced in Russia. On the other hand, in larger and more diversified cities there are more opportunities for human capital development, as well as more competition between employees and firms, all this leading to better skills, so wages are higher. Moreover, people with higher human capital do migrate to larger cities. For Russia negative dependence between wages and investments was not found in the studies considering regions. In our research we do not find negative dependence between wages and foreign firms' location choice for cities.

As for resource seeking investments, natural resources endowment (fuel) proved to be an important factor in explaining attraction of region for FDI (Castiglione et al., 2012).

Researchers find that prior domestic investment experience positively affects FDI, private investment serving a measure of economic and institutional development. For transitional economies, including Russia, the level of human capital, institutions and efficient infrastructure (measured as telephone networks, public spending on roads, road density etc.) were found to be important determinants of FDI (Castiglione et al., 2012).

Ledyaeva (2009) studies the determinants of FDI inflow into Russian regions during 1995-2006. She finds that significant determinants of FDI for Russia are presence of oil and gas, agglomeration advantages, the level of industrial development, market size and large cities advantages, benefits of transit region and protection of investors by the local legislation. High and growing regional FDI concentration in Russian economy is revealed.

Ledyaeva (2010) analyzes location choices of Chinese, Japanese and American firms in Russia using firm-level data. She finds that the important determinants of FDI location in a region are market potential, transport infrastructure, democracy development and agglomeration effects. Besides, she finds that for the firms from the same country positive agglomeration effects prevail, while for the firms from different countries competition effects dominate (Ledyaeva et al., 2010). Legislative risk, institutional potential, democracy development and bureaucracy level in the regions are taken into account. (Ledyaeva et al., 2010) confirms the results of previous research that the majority of FDI from the USA, Japan and China to Russia are market seeking. Concerning institutions, she concludes that it is institutional distance between the home and host country rather than institutional quality that

matters; only democracy in the regions positively affects FDI location from all three countries. Ledyaeva (2010) studies agglomeration effects generated by the firms working in the same industry, i.e. localization economies. Our contribution is to distinguish between localization and diversity factors; we also compare location determinants for national and foreign firms. Determinants of location choice are measured on the city level (agglomeration effects, HMP, and wage).

Using conditional logit model and data on the Russian firms with FDI and regional data for the period of 2000-2009, Gonchar and Marek (2013) find that geographical pattern of FDI location choice in Russia is determined both by natural resources availability and market-related benefits. They also find that agglomeration effects exist both for the firms of the same industry and of different industries. Another conclusion is that FDI into service sector are co-located with resource oriented FDI.

Castiglione et al. (2012) study the factors attracting FDI into the Russian regions, based on data for 79 regions for the period of transition, 1996 – 2006. They analyze cross sectional and panel data applying OLS, Prais-Winsten estimation, and GMM. They use five groups of explanatory variables: classical demand factors (gross regional product per capita to reflect market size, population to measure absolute size of the regions, distance from Moscow), geographical (coast), capital (private investments and human capital), institutional environment (social stability reflected by the percent of the Russian population, and risk captured with the index created by the Analytical agency 'Expert') and infrastructure (rail road density), and availability of fuel (oil). The dependent variable is logarithm of the sum of FDI in the region. The study showed that very uneven investment distribution is affected both by classical demand factors and by specific economic and socio-institutional regional characteristics. Castiglione et al. (2012) also conclude that FDI distribution during the second stage of transition period, 1996-2001, affected the FDI inflow in 2002-2006.

Location choice of enterprises is interrelated with economic growth. In the *Russian Economic Report* (2007) the following factors of regional growth in 1999-2004 are summarized: investment, human capital, natural resources, ports, and climate. The impact of automobile and railroad infrastructure on regional economic growth was not clear according to the *Russian Economic Report* (2007). In our study the impact of transport infrastructure on enterprise location choice for city was not clear too. In the *Russian Economic Report* (2007) urbanization measured as the size of the largest city in a region was among the most important variables explaining regional growth; proximity to the large and growing regions positively affects regional economic growth. In our study it is diversity index and HMP of the surrounding cities that affect enterprise location choice for city.

Ahrend (2008) studies the impact of a number of factors on the economic performance in Russian regions, using extreme bound analysis. He concludes that after 1998 crisis some initial conditions, such as hydrocarbon wealth or favourable geographic location continue to affect growth, but political and economic reforms have the most significant impact on economic performance.

Drapkin et al. (2011a, 2011b), based on panel data for 1999-2008, reveal that firstly, market potential and regional infrastructure positively affect revenue of foreign firms in the Russian regions. Secondly, wage level is positively correlated with foreign firms' revenue. Thirdly, regional specialization on a certain industry positively affects revenue. They found that trade openness does not have a significant impact on foreign enterprise performance in Russia. In our research location determinants both for foreign and national firms are analyzed.

Mariev and Nesterova (2011) point out the following factors of regional attractiveness for FDI: economic (GRP, wage level, education level), infrastructure development (length of roads, city population share), openness (external trade development, lagged FDI), geographical factors (climate, distance to the capital, sea port), regional economic policy (investment ranking, internal private investment), institutional (crime rate).

Concerning motivation of investors, there is literature studying choice between different types of foreign economic activity, such as export vs. 'horizontal' FDI (Helpman et al., 2004), or investment at home vs. investment abroad (Mayer et al., 2010).

According to Helpman et al. (2004), the firm's choice between exporting and FDI is associated with proximity-concentration trade off: trade costs vs. expenses to maintain several production capacities. Specific feature of exporting is lower fixed costs, while that of FDI is lower variable costs (exporting is associated with trade costs, which are variable costs). In the model developed by Helpman et al. (2004), firms with high enough productivity enter domestic market. Then if a firm's productivity is sufficiently high, it can export, and firms with the highest productivity can implement FDI. So, in line with the results of Melitz (2003), only the most productive firms serve the foreign market. Helpman et al. (2004) also find that the industries with higher firm heterogeneity carry out more FDI sales relative to export. Helpman et al. (2004) confirm their theoretical findings with the US data, mainly for 1994.

Mayer et al. (2010) explain the choice between domestic and foreign investments, in order to address the concern about off-shoring. Unlike Helpman et al. (2004), they analyze firms' decisions where to create a plant, conditional on their decision to start production. They use firm level data on French investments, in France and abroad, for 1992-2002 and apply conditional logit model. The results show that French firms over-invest in France as they benefit from agglomeration externalities generated by affiliates already working there.

Another result is that firms with higher productivity and better developed financial internationalization participate in FDI. More details on theoretical and empirical models with heterogeneous firms are provided in literature review (Chapter 1).

To summarize, the determinants of FDI location considered in the literature are the following: costs and quality of the resources (labor costs and human capital quality or labor productivity; transport costs), agglomeration effects, market potential (reflected by gross regional product, gross regional product per capita, the distance to the other markets taken into account, and by population), market growth, infrastructure, government policies, institutional factors (for example, informal institutions reflected by the corruption level) and political stability; openness, pool of technical expertise; unemployment, poverty rate (Boudier-Bensebaa, 2005; Ledyaeva et al., 2010). Market size, population, transport costs and distance are 'classical' determinants of FDI location; for transition countries with weak institutional environment, social and institutional characteristics are important (Castiglione et al., 2012). We analyze location choice of national and foreign firms for a city, taking into consideration some aspects of regional infrastructure and institutions.

5.3. Theoretical background

Enterprise location decisions are assumed to depend mainly on demand (HMP) and on agglomeration level. Specific features of national and foreign firms' location decisions are studied; institutional factors and infrastructure are taken into account. Estimation is based on the profit equation presented in the paper by Head and Mayer (2004) and in the book by (Combes et al., 2008a). Following Head and Mayer (2004), the expected profits of a firm in each of the possible locations is analyzed to predict the probability that a firm would invest in a certain location.

Equilibrium operating profits received in market *s* are:

$$\pi_{rs}^{*} = \left(p_r^{*} - m_r \right) \tau_{rs} q_{rs}^{*} = m_r \frac{\tau_{rs} q_{rs}^{*}}{\sigma - 1},$$

where *r, s* – regions; p_r - mill price set by a firm located in r; m_r - marginal production cost; q_{rs} - the quantity sold by the firm from region *r* on the market *s*; τ_{rs} - the iceberg-type trade cost from *r* to *s*; σ - elasticity of substitution between two varieties of good. The equilibrium price is defined by the following expression: $p_{rs}^{*} = \tau_{rs} p_r^{*} = \tau_{rs} m_r \frac{\sigma}{\sigma - 1}$.

The quantity is derived based on CES demand function, and the following expression is received (in the short run, with exogenous number of firms): $q_{rs}^{*} = \left(p_r^{*} \tau_{rs} \right)^{-\sigma} \mu_s Y_s P_s^{\sigma - 1}$,

where Y_s is the income in region s, and μ_s is the share of the good in consumption in the region s; the price index is defined the following way: $P_s = \left[\sum_{r=1}^{R} n_r \left(p_r^* \tau_{rs} \right)^{-(\sigma-1)} \right]^{-1/(\sigma-1)}$.

Total profit equation for a firm located in region *r* is the following:

$$\pi_r^* = \sum_s \pi_{rs}^* - F_r = cm_r^{-(\sigma-1)} RMP_r - F_r ,$$

where $c = \frac{\sigma^{-\sigma}}{(\sigma-1)^{-(\sigma-1)}}$; F_r - fixed cost; $RMP_r \equiv \sum_s \phi_{rs} \mu_s Y_s P_s^{\sigma-1}$ - real market potential (Harris, 1954) with $\phi_{rs} = \tau_{rs}^{-(\sigma-1)}$. Based on gravity model approach, market potential measures accessibility of various regions from the market r: $MP_r \equiv \sum_s \frac{Y_s}{d_{rs}}$, where d_{rs} is the distance between r and s, so that with this approach we receive inverse distance weighted sum of incomes. If we considered regions, it could be gross regional products. In case of cities it is total revenue of enterprises located in a city. In the next section specification of location choice model for empirical analysis is discussed.

5.4. Empirical background: specification of location choice model

It is assumed that firms choose locations where they can earn the highest profits. Conditional logit model is used to estimate the parameters of the profit equation, i.e. the location choice of enterprises (MacFadden, 1974, 1978; Head and Mayer, 2004; Combes et al., 2008a, ch. 12). The dependent variable is location choice for a city. The model is estimated for the national firms and for the foreign firms.

Enterprises make their decisions in sequential manner, and some market congestion arises. Firms are attracted to the region with the highest profits. It leads to increase in competition in this region and to decrease in market potential for each firm. On the other hand, demand increases leading to increase in the market potential. If real market potential decreases below a certain point, firms start choosing the other regions.

Profits are analyzed based on the following equation:

$$U_r = \ln A_r + \frac{1}{\sigma-1} \ln RMP_r - \ln CP_r ,$$

where A_r - total factor productivity in the region r; RMP_r - real market potential; CP_r - cost price for each firm, which accounts for wage and composite input price. There is also a random component specific to firm-region pairs, ε_r. Its distribution function determines functional form of probability of location choice. With firm-specific random component:

$$\overline{U}_r = \ln A_r + \frac{1}{\sigma - 1} \ln RMP_r - \ln CP_r + \varepsilon_r .$$

The probability that an enterprise will chose region r is reflected in the expression: $\beta x_{ir} + \varepsilon_{ir}$.

It is assumed that ε_r follows the Gumbel law, where cumulative distribution function is $F(\varepsilon_r) = \exp(-\exp(-\varepsilon_r))$. Then the probability of location in region *r* can be expressed by the conditional logit model (Wooldridge 2002; Ledyaeva et al., 2010):

$$P_{ir} = P(y_i = r \mid x_i) = P\left(\beta x_{ir} + \varepsilon_{ir} \geq \max_{s \in R_i, s \neq r} (\beta x_{is} + \varepsilon_{is})\right) = \frac{\exp(\beta x_{ir})}{\sum_s \exp(\beta x_{is})} = \frac{\exp U_r}{\sum_s \exp U_s}$$

During the time of Soviet Union enterprises were located in the cities and regions not according to the market economy laws but in line with needs of the centrally planned economic system. Moreover, the 'young' enterprises, registered after privatization ended in 1995, were largely based on the already existing enterprises' assets. They are likely to be located in the same territories as the ones assigned to these territories by central planning system. Therefore enterprises, particularly the national ones are not likely to consider a possibility of location in some city of Russia, i.e. to literally have a location choice. We rather analyze formation and development of business activity in a city compared to the other cities.

However, foreign investors do have a location choice; their location pattern differs from that of the national enterprises. Therefore in their case conclusions about the actual location choice can be made. The same time, it should be mentioned that from the available data the country of origin and the share of a foreign investor in the enterprise are not known. There are 172 cities, for which data on city average wages is present, i.e. 172 choices for a city that an enterprise can make.

The logit model makes possible estimation of parameters of the profit equation U_r using the maximum likelihood method (Combes et al., 2008a). The observations of the firms' location choices and regional characteristics allow sorting regions based on their potential profits and understanding the effect of the variables in the equation U_r. The dependent variable of the conditional logit model, the city chosen by each investor, depends on profit received in a region. Therefore explanatory variables are those affecting profit by affecting revenue (such as the access to demand) and production costs (such as wages).

The logit model is estimated for all enterprises, as well as for national and foreign enterprises separately in order to test the hypothesis that location (or formation and development) incentives for the national and foreign firms differ. Analysis is carried out for groups of industries, as agglomeration effects and HMP vary across industries (Head Mayer,

2004). The differences in location factors between various types of cities and firms (private and state enterprises, 'young' and 'old' enterprises) are studied. Agglomeration indices (localization, diversity and urbanization) are calculated based on the approach developed by Martin et al. (2011) and Vorobyev et al. (2010) and using data on enterprise revenue.

Total factor productivity is a function of localization, diversity and urbanization levels in a city, i.e. the variables reflecting agglomeration externalities generated by the other firms, both national and foreign (Bode et al., 2009). It is also a function of location and technological advantages: availability of resources, including human capital; enterprises' technological innovations, regional transport infrastructure and business climate. The type of a city is also assumed to affect a firm's total factor productivity: location of a firm in an agglomeration center, a city belonging to an agglomeration, a monotown. These factors can be interpreted as those forming a comparative advantage of a city.

Behind the demand aspect there is gravity model, stating the dependence of economic activity between regions on their market sizes; it is particularly relevant for the case of foreign enterprises location choice (Leyaeva, 2010). Ledyaeva (2010) applies principal component analysis to define market potential based on gross regional product (GRP), total population and population density in a region and finds that GRP has the largest impact on location choice. However, as we analyze city level location choice, total revenue of firms located in a city is used to account for home market potential (market size), reflecting industrial demand and demand for consumer goods. Market potential is weighted by distance between cities and is calculated based on Harris (1954) approach as it is done in the paper by Head and Mayer (2004). Ledyaeva (2009) uses the following approach:

$$w\left(d_{k,n}\right) = \frac{\min\_ d_{k,n}}{d_{k,n}} \forall k \neq n, \text{ (A)},$$

where $d_{k,n}$ - the distance between regions *k* and *n* taken as the distance between their capitals; $\min\_ d_{k,n}$ - the minimum distance in the sample.

We use the approach (B) below, not the formula (A) above, to construct weighs:

$$w\left(d_{k,n}\right) = \frac{1}{d_{k,n}} \forall k \neq n \text{ (B)},$$

The following model is estimated to determine firm's location decision for a city:

$$y_{ijt}^{z} = \beta_1 \ln A_{ijt}^{z} + \beta_2 \ln HMP_{it}^{z} - \beta_3 \ln cp_{ijt}^{z} + \varepsilon_{ijt}^{z} \quad (5.1)$$

where y_{ijt}^{z}, the dependent variable, is enterprise location choice for a city; A_{ijt}^{z} - total factor productivity; HMP_{it}^{z} - home market potential; cp_{ijt}^{z} - cost price for each firm, which accounts for wage in our model, composite input price not being known. Index *z* denotes a city; *r* – a

region; j – industry; i – firm; t – time period. Total factor productivity, A_{ijt}^{z}, depends on the following factors (definitions are in Table 5.1 below):

$$A_{ijt}^{z} = f(loc_{ijt}^{z}, div_{jt}^{z}, urb_{jt}^{z}, transp\_infrastructure_{jt}^{r}; busnenvrisk_{jt}^{r}) \quad (5.2)$$

Overall, independent variables are those affecting profit through the impact on revenue and costs: agglomeration economies, potential demand, resource availability and costs, location advantages and infrastructure. To reflect labor costs and quality, wages are used. An alternative variable to reflect labour quality would be educational level of labor force - the share of workforce with higher education or also with vocational training.

Transportation costs are reflected in distance component of HMP and by automobile and railroad density. Following Ledyaeva (2010) an institutional variable is included into the model – regional business environment risks. While Ledyaeva (2010) analyzes foreign direct investment, here the analysis is extended to national enterprises too. In the next section, estimation procedure is discussed.

5.5. Econometric model

The dependent variable is enterprise choice for a city. Agglomeration levels, home market potential, costs of doing business, transport infrastructure and business environment are considered as explanatory variables. These variables are presented in Table 5.1 with comments provided below. All agglomeration indices are measured for 3-digit level of OKVED classification.

Table 5.1. Definition of variables used in the econometric analysis

Variable	Definition
Agglomeration indices (city level)	
$\ln(loc)_{ti}^{jz}$	Localization coefficient (Equation (5.3))
$core_{t}^{jz}$	A share of an industry j in the total revenue in a city z
$\ln(div)_{t}^{z}$	Diversity coefficient (Equation (5.4))
$\ln(\text{urb})_{t}^{jz}$	Urbanization coefficient (Equation (5.5))
Home market potential (city level): HMPcity= $\ln(urb)_{t}^{jz} + \ln(hmpcity)_{t}^{z}$, where	
$\ln(hmpcity)_{t}^{z}$	Home market potential of cities, other than a city where enterprise is located $hmpcity_{t}^{Z} = \sum_{z \neq Z} \frac{revenue_{t}^{z}}{dist_{Z,z}}$
City characteristics	
$\ln(\text{wagecity})_{t}^{z}$	Average monthly nominal wage in a city, payroll, roubles

Variable	Definition
Regional transport infrastructure	
$\ln(\text{a\_road})_t^z$	Density of auto roads in region r, end of year, km of roads per 1000 sq km of territory
$\ln(\text{rw\_road})_t^z$	Density of railroads in region r, end of year, km of roads per 1000 sq km of territory
Regional investment climate (Indices by analytical agency 'Expert')[50]	
$\ln\left(\text{expertinvptl}\right)_t^r$	Investment potential in region r at time t (Regional investment potential by the Analytical agency Expert)
$\ln(\text{busnenvrisk})_t^r$	Business environment risks, region r, time t (Regional investment risk by the Analytical agency Expert)

Following Martin et al. (2011) and Vorobyev et al. (2010), we apply the following indicator for measuring *localization*:

$$\ln(loc)_{it}^{jz} = \ln(revenue_t^{jz} - revenue_{it}^{jz} + 1), \tag{5.3}$$

where $revenue_t^{jz}$ – the revenue of all firms belonging to an industry *j* and located in a city *z*; $revenue_{it}^{jz}$ – revenue of a firm *i* belonging to an industry *j* and located in a city *z*. This indicator, although with labor instead of revenue, was applied by Martin et al. (2011).

Concerning *diversity*, a version of index suggested by Vorobyev et al. (2010) is used. They suggest a new indicator of diversity, taking into account both variety of industries and inequality among industries in city (details are provided in Chapter 2):

$$div_t^z = \frac{\sum_{j=1}^{s}\left(\frac{revenue_t^{jz}}{revenue_t^z}\right)^{\frac{1}{s}} - 1}{s^{1-\frac{1}{s}} - 1}. \tag{5.4}$$

This coefficient takes on the value 1 if the distribution of revenue among industries is perfectly equal, and the value 0 if all revenue is concentrated in one industry. Diversity is measured on a city level. It is closely linked to the city size, i.e. city area or the amount of business activity (total revenue of firms), and to the population density. In the version of this coefficient used in our work, *s* reflects the number of industries in the country.

Urbanization coefficient measures total revenue of firms belonging to all industries in a city, except for the industry under consideration:

$$\ln\left(\text{urb}\right)_t^{jz} = \ln(revenue_t^z - revenue_t^{jz} + 1) \tag{5.5}$$

where $revenue_t^{jz}$ – revenue of all firms belonging to an industry *j* and located in a city *z;* $revenue_t^z$ - revenue of all firms in a city *z*.

Home market potential of surrounding cities is the sum of total revenues in all industries in cities divided into the distances between the city *Z* and the other cities:

[50] Details are in Appendix C 'Data sources'.

$$hmpcity_t^Z = \sum_{z \neq Z} \frac{revenue_t^z}{dist_{Z,z}} \quad (5.61)$$

where $dist_{Z,z}$ are the *physical* distances between the cities. The indices presented in equations (5.3) – (5.6) are described in more detail in Chapter 2.

The variable lwagecity can be interpreted as a proxy for human capital; it also partly captures resource availability and resource costs, reflecting comparative advantage of a city. Therefore, wage affects expected profits in two ways. On one hand, higher wages lead to higher costs reducing the expected profits. On the other hand, higher wages may reflect higher human capital level, and increase the expected profits. The variable area would be useful, as larger city area is likely to imply higher costs for a firm, including higher wages. However, this variable proved to be insignificant; its correlation with diversity index is rather high (55%).

The variables ln(a_road) and ln(rw_road) account for automobile road and railroad density, i.e. transport infrastructure in a region where a city is located. Concerning institutional factors, Ledyaeva et al. (2010) included into their model institutional potential and legislative risk measured by the Analytical agency Expert. These indices are parts of investment potential and investment risk respectively. As significance of investment potential and investment risk components varies across years according to experts' estimation, we include the overall indices of investment potential and investment risk (regional business environment risk). Investment potential was less important from the investors' point of view than investment risk during the period 1999-2008[51]. For our sample, it is also only investment risk (regional business environment risk) that proved to be significant, therefore finally we include only this variable into the model. Details on the Expert indices are in Appendix C.

Other regional factors that we tried to include into the model, proved to be insignificant: innovations index (geometric mean of expenditures for technological innovations, the number of personnel employed in research and development, percentage of organizations involved in technological innovations in total number of organizations and the number of patents), share of economic activity 'natural resource extraction' in structure of shipped products (works, services), population with income below subsistence level (percent from total population of the region), the level of registered unemployment, the level of economic activity of population (percent from total population of the region).

To test the probability of business location in a city, equation (5.1) is estimated using conditional logit model described in Section 5.3 above. All regressors are alternative specific,

[51] Ranking of regional investment attractiveness. Analytical agency 'Expert', 1996-2013 http://www.raexpert.ru/ratings/regions/ratingclass/

alternatives being the cities among which enterprise chooses. Stata command *xtlogit* with enterprise level fixed effects is applied, dimensions of panel in *xtlogit* being enterprises and cities, among which an enterprise chooses. Using the Stata command *clogit* with robust standard errors, the coefficients and their significance are very similar.

5.6. Data and descriptive statistics

Research is based on the SPARK-Interfax data augmented with data covering regional and city characteristics. SPARK-Interfax database contains enterprise level data on the organizational form, property, the year of foundation, location, revenue, labour, cost price and profit. Data for the year 2007 is used. Indexes are constructed based on 3-digit OKVED classification. The sample includes manufacturing tradable industries[52]. For the year 2007 there are 8569 firms of tradable industries located in cities with population over 100 thousand people, for which there is data on city average wage, used in estimation.

As there is a tendency for businesses that work in various regions of Russia to register in Moscow, the indicators of localization and diversity in Moscow are probably exaggerated. The cities Moscow and St. Petersburg (which are federal cities), are excluded from the sample. However, the results with and without Moscow and St. Petersburg are compared below. The other subjects of Federation excluded from the sample are Chechnya Republic and 4 districts ('okrugs'). The lists of monotowns are taken from the book by Zubarevich (2010) and from the report *Monotowns in Russia* (2008). Rosstat city and regional data is used. In Tables 5.2 and 5.3 below there is summary statistics of the data used in this chapter (the sample including all cities). More details on data are provided in Chapter 2.

Table 5.2. Descriptive statistics of explanatory variables

Variable	Obs	Mean	Std. Dev.	Min	Max
divpq	949264	0.18	0.13	0.01	0.82
ldivpq	949264	-2.03	0.83	-4.93	-0.20
lloc	949264	12.34	4.85	0	23.96
lloc2	949264	175.69	110.94	0	574.03
lhmecity	949264	23.42	0.49	21.98	24.38
lurbpq	947210	23.80	1.38	12.58	26.75
lwagecity	949264	9.50	0.41	8.60	12.05

[52] SPARK: OKVED 10 (CA) – 37(DN), 40(E), 41(E) – i.e.whole E, 45(whole F), 60(part of I), 65(part of J), K and 'high tech' industries

Variable	Obs	Mean	Std. Dev.	Min	Max
la_road	945788	4.68	0.96	1.69	6.28
lrw_road	942786	5.01	0.86	2.20	6.37
lbusnenvrisk	947210	3.16	0.91	0	4.48
private	947210	0.77	0.42	0	1
state	947210	0.35	0.48	0	1
fdi	947210	0.05	0.22	0	1
new	947210	0.38	0.48	0	1
productivity	839018	1578490	4.63e+07	2.49	4.02e+09
highprod	949264	0.21	0.41	0	1
highprodmed	949264	0.55	0.50	0	1

Among the firms in the sample, diversity varies from 0.007 to 0.82, mean of diversity being 0.175 (the diversity index is constructed so that it varies from 0 to 1). Logarithm of localization index varies from 0 (if there are no other enterprises of a certain industry in a city) to 23.95. Logarithm of city HMP arising from the other cities (weighed by distances) varies from around 22 to 24.38. Logarithm of urbanization index varies from 12.6 to 26.7.

The share of firms with private ownership in the sample is 0.77; the share of firms with state ownership is 0.35, and with foreign ownership – 0.053. Firms founded after the year 1995 account for 0.375 of all firms. Productivity of 0.21 firms is higher than mean value of productivity in tradable industries. Mean and median are calculated using the sample of tradable firms for the years 1999-2008; the mean productivity for all tradable firms is 948940.3 roubles per worker; the median productivity is 253650.2 roubles per worker.

Table 5.3. Descriptive statistics: types of cities

Variable	Obs	Mean	Std,Dev,	Min	Max
agglomeration center	2003366	0.425	0.494	0	1
within agglomeration	2003366	0.043	0.202	0	1
monotown[53] within agglomeration	2003366	0.002	0.046	0	1
monotown outside agglomeration	2003366	0.053	0.224	0	1
monotown according to the definition by N.Zubarevich (2010)	2003366	0.015	0.122	0	1
other cities	2003366	0.480	0.500	0	1

In the sample considered in this study, 0.42 enterprises are located in the agglomeration centers; 0.04 enterprises – in the other cities within agglomeration. The share

[53] *Monotown* according to the definition in *Monotowns in Russia* (Monogoroda Rossii, 2008)

of enterprises located in monotowns is around 0.06, monotowns within agglomerations accounting for only 0.002 of firms.

Correlation between explanatory variables for the sample for year 2007 considered in this chapter, is presented in Appendix A, for the samples with and without Moscow and St. Petersburg (in Tables A.5 and A.6 respectively). The relations between the variables are similar to those for the years 1999-2008. The next section is devoted to the estimation results.

5.7. Estimation results

In Table 5.4 the results for equation (5.1) are presented (variables affecting Total Factor Productivity are reflected in equation (5.2)). Estimation is carried out for all enterprises producing tradable goods, as well as for the national and foreign firms separately. Moscow and St. Petersburg are excluded, except for Model (5) in Table 5.4 below. The majority of coefficients are robust to presence of Moscow and St. Petersburg.

Table 5.4. Enterprise location choice for city; national and foreign firms
Dependent variable: enterprise location choice for city

	(1) All tradables	(2) National firms	(3) Foreign firms	(4) Foreign firms	(5) Foreign firms (with Moscow and St.P.)
ldivpq	1.116***	1.115***	1.137***	1.137***	1.238***
	(30.63)	(29.80)	(7.09)	(7.09)	(8.25)
core				13.01*	14.59**
				(1.79)	(2.11)
core2				-17.15**	-21.55***
				(-2.06)	(-2.71)
lloc	0.243***	0.247***	0.170***	0.170***	-0.404***
	(14.37)	(14.13)	(2.58)	(2.59)	(-11.86)
lloc2	-0.00759***	-0.00770***	-0.00542**	-0.00543**	0.0198***
	(-12.30)	(-12.07)	(-2.21)	(-2.22)	(15.48)
lhmecity	0.201***	0.181***	0.570***	0.572***	0.396***
	(6.14)	(5.38)	(3.85)	(3.86)	(2.83)
lurbpq	0.0285	0.0441	-0.194	-0.188	-0.00848
	(0.40)	(0.60)	(-0.67)	(-0.65)	(-0.03)
lwagecity	-0.291***	-0.322***	0.150	0.149	0.512***
	(-5.76)	(-6.15)	(0.82)	(0.82)	(3.08)
la_road	-0.132	-0.0599	-1.430	-1.454	-0.785
	(-0.39)	(-0.17)	(-1.04)	(-1.06)	(-0.71)

lrw_road	2.449	2.335	6.232	6.290	10.70
	(1.17)	(1.09)	(0.60)	(0.60)	(0.99)
lbusnenv risk	-0.0492	-0.0131	-0.471*	-0.466*	0.000938
	(-0.61)	(-0.16)	(-1.77)	(-1.74)	(0.01)
N	940100 (158cities; 5950 firms)	890172 (158 cities; 5634 firms)	49928 (158 cities; 316 firms)	49928 (158 cities; 316 firms)	120396 (158 cities; 762 firms)
Log likelihood	-27925.24	-26431.612	-1482.951	-1481.2635	-2249.0194

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Agglomeration effects proved to be significant both for national and foreign firms. Coefficient of diversity for the national firms is 1.115, and for the foreign enterprises 1.137. When Moscow and St. Petersburg are included into the sample, the diversity coefficient for the foreign enterprises very slightly increases, becoming 1.238.

Localization economies seem to have inverted U shape and to be relatively higher for national enterprises. Maximum of localization economies is reached at low level: exp(0.243/(0.00759*2))=exp(16.008)=8.96 mn roubles for all tradables, more precisely, exp(16.039)=9.2 mn for the national firms and exp(15.68)=6.47 for the foreign firms. However, localization economies stay positive up to exp(0.243/(0.00759))=exp(32.016)=80 bn roubles for all tradables, exp(32.078)=85 bn roubles for the national firms, and exp(31.36)=42 bn roubles for the foreign firms. Up to the optimum localization level, localization economies are reflected by coefficient 0.243; the coefficient for the national firms being 0.247, and for the foreign firms 0.17. After the optimum is reached, probability of location starts slightly decreasing relative to the optimum, but localization effects stay positive. The coefficients reflecting this decrease are 0.00759 for all enterprises producing tradable goods, 0.0077 for the national firms, and 0.00542 for the foreign firms. These maximum numbers are not very high relative to the large company revenue in Russia[54]. The same time, mean logarithm of localization is 12.34; maximum is 23.96.

An 'inverted U shape' of agglomeration effects in a city is confirmed for localization economies. This pattern might arise because after a certain point of localization is reached, competition and congestion costs outweigh positive effects. Among the consequences are additional expenditures faced by the firms. When Moscow and St. Petersburg are included, the shape of localization economies changes into the U-shape, which is counter-intuitive. Urbanization coefficient is insignificant in all versions of the model above (Models (1) – (5)).

[54] Ranking of the largest companies in Russia in 2012 based on sales volume. Expert-400. http://expert.ru/ratings/expert400-2012/?n=897345

HMP is reflected by two components: market potential of a city and market potential of the other cities in the country taking the distance into account. Market potential of a city is reflected by the urbanization and diversity coefficients; diversity coefficient is positive and significant. The coefficient of market potential of the other cities is positive and significant; it is 0.181 for the national enterprises and 0.57 for the foreign enterprises. When Moscow and St. Petersburg are included into the sample, the coefficient becomes 0.396.

As it was mentioned above, wages effect is controversial, because wages reflect both human capital quality and costs faced by the firms. It seems that for the national enterprises costs outweigh the impact of human capital quality, as the coefficient is significant and negative, -0.322. For the foreign enterprises the coefficient is not significant. When Moscow and St. Petersburg are included into the sample, the coefficient becomes positive and significant for the foreign firms at 0.01 significance level, with value 0.512. It might mean that for the foreign firms registered in Moscow and St. Petersburg high quality of human capital reflected by higher wages is relatively more important than minimizing costs by paying lower wages. However, in a similar sample in the paper by Gonchar and Marek (2013) wages are insignificant (with negative sign) for manufacturing.

The variables accounting for railroad and automobile road infrastructure turned out to be insignificant. Business environment risk (based on the index of investment risks by the Analytical agency 'Expert') proved to be significant and negative for the foreign enterprises, being insignificant for the national enterprises. For the foreign enterprises the coefficient of business environment risk is 0.471.

Concerning the MNCs strategies (Dunning, 1993), a variable can be assigned to each incentive. Firstly, home market potential would account for market seeking strategy. The table above leads to the conclusion that foreign firms pursue market seeking strategies, HMP and diversity variables being positive and significant. This finding is in line with conclusions of the existing research on Russia (Gonchar and Marek, 2013). Secondly, wages reflect efficiency seeking strategy. Unlike national firms, foreign firms do not explicitly tend to locate in cities with lower wages, i.e. do not seem to pursue efficiency seeking strategies.

'Share of economic activity "natural resource extraction" in structure of shipped products (works, services)' is associated with natural resources seeking strategy. This variable turned out to be insignificant. As the sample contains only manufacturing enterprises, and natural resources seeking strategies are probably more likely to be pursued by the enterprises in the resource extraction sector, this strategy could not be fully explored here. However, Gonchar and Marek (2013) find that natural resources availability is important in

geographical pattern of FDI location choice in Russia. Ahrend (2006) emphasizes that fuel sector had the major role in driving economic growth, especially in 2001-2004.

Finally, the variable reflecting regional enterprises innovations level would account for strategic asset seeking. However, it turned out to be insignificant, and was not included into the model presented here. The same time, it can be assumed that there are more innovations in diversified places and in places with favourable business climate. The variables *diversity* and *regional business environment risks* are significant, diversity being highly significant.

The results in Table 5.4 stay robust when the variables *core* and *core2* are added (Models 3 and 4). The variable *core* measured at the city level is similar to the variable *special* in the paper by Gonchar and Marek (2013)[55]. The results for the manufacturing sectors are similar. Gonchar and Marek (2013) received coefficient 14.56; the null hypothesis that the coefficient is zero was rejected at 0.01 significance level. The dependence between the choice and *special* has linear form. In our case, the dependence is non-linear; the coefficient is 13.01, at 0.1 significance level. When Moscow and St. Petersburg are included into the sample, the coefficient becomes 14.59; the null hypothesis that the coefficient is zero can be rejected at 0.05 significance level.

In Tables 5.5 - 5.7 below, the location incentives are explored for different types of industries. All enterprises producing tradable goods are considered in Table 5.5; the results for the national firms are presented in Table 5.6, and the results for the foreign firms are presented in Table 5.7.

Table 5.5. Enterprise location choice for city, by class of industry; all tradables
Dependent variable: enterprise location choice for city

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
ldivpq	1.116***	1.348***	1.377***	1.327***	1.459***	1.071***
	(30.63)	(11.85)	(17.76)	(12.42)	(4.75)	(15.56)
lloc	0.243***	0.226***	0.173***	0.225***	0.219*	0.498***
	(14.37)	(4.82)	(5.85)	(5.08)	(1.79)	(11.42)
lloc2	-0.00759***	-0.00892***	-0.00668***	-0.00724***	-0.00989**	-0.0142***
	(-12.30)	(-5.04)	(-5.97)	(-4.42)	(-2.10)	(-9.73)
lhme city	0.201***	0.220**	0.289***	0.489***	0.100	0.258***
	(6.14)	(2.34)	(4.49)	(5.31)	(0.41)	(4.06)

[55] *special* is the share of employees in the sector of investment *s* compared to the total regional employment figure (Gonchar and Marek, 2013).

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
lurbpq	0.0285	-0.0827	0.101	0.220	0.0580	-0.0938
	(0.40)	(-0.39)	(0.73)	(1.22)	(0.13)	(-0.66)
lwage city	-0.291***	-0.111	-0.339***	-0.616***	-0.240	-0.125
	(-5.76)	(-0.79)	(-3.37)	(-4.14)	(-0.63)	(-1.31)
la_road	-0.132	-0.0391	-0.168	-0.280	-2.815**	0.185
	(-0.39)	(-0.04)	(-0.27)	(-0.33)	(-2.07)	(0.27)
lrw_ road	2.449	-2.764	4.169	3.950	4.368	1.197
	(1.17)	(-0.51)	(1.01)	(0.60)	(0.25)	(0.30)
lbusn envrisk	-0.0492	0.00253	-0.177	0.234	-0.244	-0.166
	(-0.61)	(0.01)	(-1.15)	(1.11)	(-0.46)	(-1.12)
N	940100	113918	259120	139198	16590	255486
Log likelihood	-27925.2	-3393.8	-7676.3	-4063.4	-494.4	-7388.4

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5.6. Enterprise location choice for city, by class of industry; national firms
Dependent variable: enterprise location choice for city

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
ldivpq	1.115***	1.320***	1.391***	1.328***	1.433***	1.084***
	(29.80)	(11.41)	(17.21)	(12.11)	(4.50)	(15.51)
lloc	0.247***	0.214***	0.188***	0.214***	0.191	0.511***
	(14.13)	(4.50)	(6.00)	(4.77)	(1.55)	(11.47)
lloc2	-0.0077***	-0.00830***	-0.0071***	-0.00707***	-0.00897*	-0.0147***
	(-12.07)	(-4.64)	(-6.02)	(-4.24)	(-1.87)	(-9.85)
lhme city	0.181***	0.196**	0.256***	0.452***	0.0618	0.250***
	(5.38)	(2.03)	(3.86)	(4.80)	(0.24)	(3.90)
lurbpq	0.0441	-0.0745	0.0956	0.245	0.0879	-0.0640
	(0.60)	(-0.34)	(0.67)	(1.34)	(0.19)	(-0.45)
lwage city	-0.322***	-0.107	-0.403***	-0.606***	-0.220	-0.115
	(-6.15)	(-0.74)	(-3.79)	(-3.98)	(-0.56)	(-1.19)
la_road	-0.0599	0.274	-0.120	-0.271	-3.645***	0.102
	(-0.17)	(0.28)	(-0.18)	(-0.31)	(-2.63)	(0.15)
lrw_	2.335	-2.796	4.013	3.852	5.575	0.508

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
road						
	(1.09)	(-0.51)	(0.95)	(0.58)	(0.33)	(0.13)
lbusn envrisk	-0.0131	0.0804	-0.125	0.243	-0.306	-0.129
	(-0.16)	(0.34)	(-0.78)	(1.13)	(-0.57)	(-0.84)
N	890172	107756	242214	130982	15010	250746
Log likelihood	-26431.6	-3213.1	-7155.7	-3835.3	-448.3	-7246.8

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5.7. Enterprise location choice for city, by class of industry; foreign firms
Dependent variable: enterprise location choice for city

	(1) All tradables	(2) Traditional	(3) Basic materials	(4) Basic machinery	(5) Integrating machinery	(6) Science based
ldivpq	1.137***	2.054***	1.181***	1.355***	1.917	0.584
	(7.09)	(3.52)	(4.26)	(2.85)	(1.63)	(1.51)
lloc	0.170***	0.604**	0.0126	0.660**	0.717	0.0534
	(2.58)	(2.03)	(0.14)	(2.18)	(1.06)	(0.27)
lloc2	-0.00542**	-0.0268**	-0.00181	-0.0178*	-0.0266	0.00314
	(-2.21)	(-2.34)	(-0.50)	(-1.80)	(-1.12)	(0.44)
lhme city	0.570***	0.793*	0.723***	1.179***	0.523	0.840
	(3.85)	(1.73)	(2.78)	(2.67)	(0.60)	(1.64)
lurbpq	-0.194	-0.249	0.142	-0.384	-0.896	-1.010
	(-0.67)	(-0.31)	(0.29)	(-0.38)	(-0.33)	(-1.48)
lwage city	0.150	-0.166	0.244	-0.661	-0.242	-0.600
	(0.82)	(-0.26)	(0.85)	(-0.96)	(-0.17)	(-0.81)
la_road	-1.430	-7.336**	-0.709	-0.637	3.919	2.423
	(-1.04)	(-2.33)	(-0.29)	(-0.14)	(0.32)	(0.53)
lrw_ road	6.232	2.759	8.358	3.586	-11.85	67.76***
	(0.60)	(0.14)	(0.40)	(0.09)	(-0.09)	(2.58)
lbusn envrisk	-0.471*	-1.022	-0.703	0.0442	-0.0697	-0.362
	(-1.77)	(-1.50)	(-1.54)	(0.04)	(-0.03)	(-0.71)
N	49928	6162	16906	8216	1580	4740
Log likelihood	-1483.0	-174.8	-511.2	-222.4	-44.43	-135.1

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Model (5) in Table 5.7 above is not statistically significant according to the LR chi2 test (probability=0.1918). The other models are statistically significant. Diversity is assumed to be relatively more important for science based industries. However, Table 5.5 shows that it is integrating machinery industries that benefit relatively more from diversity, followed by basic materials, traditional industries, basic machinery and science based industries, if all firms are considered. Table 5.6 shows that for the national firms, diversity effect is the highest for integrating machinery again (1.433), followed by basic materials (1.391), basic machinery (1.328), traditional (1.320) and science based (1.084). According to Table 5.7, for the foreign enterprises diversity effect is the highest for traditional industries (2.054), followed by basic machinery (1.355) and basic materials (1.181). For science based industries the coefficient is insignificant, possibly due to lack of observations; integrating machinery is not considered because Model (5) is not statistically significant as it was mentioned above.

Localization effect is the highest for science based industries, followed by traditional and basic machinery, integrating machinery, and basic materials, if all enterprises are considered. For the national enterprises the sequence is the same: localization effect is the highest for science based industries (0.511), followed by traditional industries and basic machinery (0.214), and basic materials (0.118). For integrating machinery the coefficient is insignificant; however, for this group the number of observations is much lower compared to the others. For the foreign firms, localization effect is the highest for basic machinery industries (0.66), followed by traditional industries (0.604); for the other industries it is insignificant, possibly due to lack of observations. It turned out that localization effects are higher for the foreign firms than for the national firms, if to look at the groups of industries separately. Localization effect keeps the inverted U shape for all groups of industries.

For all firms producing tradable goods, HMP arising from the other cities is the largest for basic machinery industries, followed by basic materials industries, science based and traditional industries; for integrating machinery it is not significant. For the national enterprises the sequence is the same: the effect is the largest for basic machinery industries (0.452), followed by basic materials industries (0.256), science based (0.25) and traditional industries (0.196). For the foreign enterprises the effect is the highest for basic machinery industries (1.179), followed by traditional (0.793) and basic materials (0.723), and is insignificant for integrating machinery and science based industries.

Overall, in line with our hypothesis, foreign enterprises, having more opportunities to choose location, benefit more than the national ones from agglomeration (particularly, localization) and HMP. This difference might also occur due to differences in management

practices (Vakhitov, 2008). The results for road infrastructure are ambiguous. The effect of wages is in line with the results for the whole samples of all tradables, national and foreign firms. The result for business environment risks does not hold for the groups of industries.

In Table 5.8 below, enterprise location choice is analyzed for enterprises with relatively higher and relatively lower productivity levels.

Table 5.8. Enterprise location choice for city, by productivity; all tradables
Dependent variable: enterprise location choice for city

	(1) All tradables	(2) Firms with productivity above the mean[56]	(3) Firms with productivity below the mean
ldivpq	1.116*** (30.63)	0.930*** (12.91)	1.182*** (27.85)
lloc	0.243*** (14.37)	0.246*** (6.69)	0.240*** (12.58)
lloc2	-0.00759*** (-12.30)	-0.00692*** (-5.25)	-0.00774*** (-11.05)
lhmecity	0.201*** (6.14)	0.196*** (2.80)	0.203*** (5.46)
lurbpq	0.0285 (0.40)	0.127 (0.82)	-0.000236 (-0.00)
lwagecity	-0.291*** (-5.76)	0.0412 (0.43)	-0.401*** (-6.77)
la_road	-0.132 (-0.39)	0.0442 (0.06)	-0.166 (-0.44)
lrw_road	2.449 (1.17)	2.556 (0.52)	2.390 (1.02)
lbusnenvrisk	-0.0492 (-0.61)	0.105 (0.59)	-0.0952 (-1.04)
N	940100	184492	729465
Log likelihood	-27925.2	-5693.2	-21826.5

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In Table 5.8, diversity turned out to be relatively more important for the firms with productivity below the mean, while localization proved to be very slightly more important for the firms with productivity above the mean. The difference in HMP is not very large for the two types of firms. Concerning wages, estimation shows that for higher productivity firms wage level is insignificant, which is probably due to the controversial effect of wage

[56] The mean productivity for all tradable firms is 948940.3 roubles of revenue per worker

discussed above. Lower productivity firms tend to locate in the cities with lower wages. The other variables are insignificant. In Table 5.9 below, enterprise location choice is analyzed for the private and state firms, and for the 'young' and 'old' firms.

Table 5.9. Enterprise location choice for city: private and state firms; 'young' and 'old' firms

Dependent variable: enterprise location choice for city

	(1) All tradables	(2) Private firms	(3) State firms	(4) 'Young' firms	(5) 'Old' firms
ldivpq	1.116***	1.088***	1.229***	0.980***	1.207***
	(30.63)	(26.39)	(19.09)	(17.62)	(25.13)
lloc	0.243***	0.229***	0.285***	0.249***	0.239***
	(14.37)	(12.07)	(9.43)	(9.01)	(11.16)
lloc2	-0.00759***	-0.00694***	-0.00967***	-0.00713***	-0.00788***
	(-12.30)	(-9.99)	(-8.84)	(-7.13)	(-10.04)
lhmecity	0.201***	0.264***	0.0864	0.257***	0.167***
	(6.14)	(6.99)	(1.58)	(4.83)	(4.01)
lurbpq	0.0285	0.0720	-0.0281	0.0756	-0.00364
	(0.40)	(0.88)	(-0.23)	(0.66)	(-0.04)
lwagecity	-0.291***	-0.367***	-0.0991	-0.230***	-0.331***
	(-5.76)	(-6.28)	(-1.20)	(-2.91)	(-5.05)
la_road	-0.132	-0.0719	0.208	-0.122	-0.135
	(-0.39)	(-0.18)	(0.36)	(-0.22)	(-0.31)
lrw_road	2.449	3.522	1.400	1.773	2.705
	(1.17)	(1.46)	(0.41)	(0.49)	(1.06)
lbusnenvrisk	-0.0492	0.00855	-0.182	0.0350	-0.101
	(-0.61)	(0.09)	(-1.31)	(0.26)	(-0.99)
N	940100	724272	323426	354078	586022
Log likelihood	-27925.2	-21523.8	-9580.6	-10552.3	-17364.6

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Tables 5.8 - 5.9 above show that less productive firms, firms with state ownership and older firms benefit from diversity more than the other ones. Localization effect is slightly larger for the firms with state ownership and for the 'young' firms. At the first glance, this result contradicts the product cycle concept (Neffke, 2009), according to which for the younger enterprises being at exploration (or search) stage diversity externalities are relatively more important, while localization externalities are more important for the older firms being at exploitation (or mass production) stage. However, after the year 1995 a lot of enterprises were founded based on the assets of the existing enterprises; therefore the difference between the 'young' and 'old' enterprises may not be so explicit.

HMP is significant for the firms with private ownership and insignificant for firms with state ownership. It is significant both for 'young' and for 'old' firms, and is larger for the 'young' firms. Private firms are interested in lower wages, while for state firms wages are insignificant. Both for 'young' and 'old' firms wage coefficient is significant and negative, being larger for the 'old' firms.

Table 5.10. Enterprise location choice for city; differences among types of cities
Dependent variable: enterprise location choice for city

	(1) All tradables	(2) National firms	(3) Foreign firms
ldivpqac	4.715***	4.798***	3.623***
	(27.79)	(27.22)	(5.99)
llocac	5.515***	5.414***	7.688***
	(25.12)	(24.31)	(6.29)
lloc2ac	-0.157***	-0.155***	-0.209***
	(-27.10)	(-26.32)	(-6.47)
lurbpqac	-0.153	-0.102	-0.524
	(-1.08)	(-0.69)	(-1.03)
lhmecityac	-0.355***	-0.344***	-0.658*
	(-4.29)	(-4.07)	(-1.68)
lwagecityac	3.040***	3.060***	2.683***
	(18.65)	(18.37)	(3.78)
lbusnenvriskac	0.148	0.250	-0.864
	(0.82)	(1.34)	(-1.46)
ldivpqwa	-0.0266	-0.0128	-0.132
	(-0.17)	(-0.08)	(-0.27)
llocwa	0.107	0.0968	0.221
	(1.23)	(1.06)	(0.74)
lloc2wa	-0.00671*	-0.00626	-0.0115
	(-1.67)	(-1.48)	(-0.85)
lurbpqwa	0.474**	0.518**	-0.106
	(2.02)	(2.14)	(-0.12)
lhmecitywa	5.132***	5.249***	4.041***
	(20.38)	(19.68)	(5.27)
lwagecitywa	0.230**	0.244**	0.0833
	(2.39)	(2.44)	(0.22)
lbusnenvriskwa	-0.188	-0.250	0.365
	(-0.63)	(-0.81)	(0.37)
ldivpqm	-0.197***	-0.224***	0.118
	(-2.79)	(-3.06)	(0.42)
llocm	0.127**	0.136**	0.0469
	(2.53)	(2.56)	(0.32)
lloc2m	-0.00409**	-0.00436**	-0.00211
	(-1.98)	(-1.99)	(-0.34)
lurbpqm	-0.294*	-0.339**	0.116
	(-1.80)	(-2.02)	(0.22)
lhmecitym	0.657***	0.668***	0.548
	(6.08)	(5.92)	(1.47)
lwagecitym	0.545***	0.541***	0.621**
	(6.61)	(6.27)	(2.21)
lbusnenvriskm	-0.392	-0.345	-0.698
	(-1.50)	(-1.26)	(-1.01)
ldivpqco	0.984***	0.978***	1.163***
	(22.43)	(21.90)	(5.03)
llocco	0.310***	0.312***	0.275**
	(12.88)	(12.62)	(2.56)

	(1) All tradables	(2) National firms	(3) Foreign firms
lloc2co	-0.00962***	-0.00961***	-0.00997**
	(-10.61)	(-10.33)	(-2.39)
lurbpqco	0.0911	0.112	-0.377
	(1.05)	(1.26)	(-0.97)
lhmecityco	-0.0111	-0.0315	0.454**
	(-0.26)	(-0.74)	(2.15)
lwagecityco	-1.851***	-1.879***	-1.273***
	(-23.83)	(-23.63)	(-3.52)
lbusnenvriskco	-0.0722	-0.0579	-0.271
	(-0.71)	(-0.55)	(-0.74)
la_road	-0.114	-0.0742	-0.626
	(-0.33)	(-0.21)	(-0.50)
lrw_road	2.455	2.410	1.717
	(1.18)	(1.13)	(0.16)
N	940100	890172	49928
Log likelihood	-25172.565	-23833.438	-1323.1745

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In Table 5.10 above, variables with 'ac' stand for the variables multiplied by the dummy variable 'agglomeration center'; 'wa' – a city within agglomeration, but not an agglomeration center; 'm' – monotown belonging to at least one of the two lists of monotowns referred to in this work; 'oc' – all other cities. Variables la_road and lrw_road accounting for transport infrastructure were not subdivided into those characterizing different types of cities because they are insignificant in all specifications above.

The results show that if choosing an agglomeration center for location, diversity is relatively more important for the national firms than for the foreign firms; for the foreign firms localization is more important. The effect of HMP proved to be negative for the choice of an agglomeration center. This effect is higher for the foreign firms. Wage effect is positive and slightly higher for the national firms.

Concerning probability of choosing a city or town within agglomerations, other than agglomeration center, diversity and localization effects are not significant; urbanization effect is significant for the national firms; HMP is significant for all firms, and is higher for the national firms; wages are significant and positive for the national firms and insignificant for the foreign firms. For the cities and towns within agglomerations, other than agglomeration centers, the number of observations is relatively low, 4.3% of all cities (Table 5.3).

Both for the foreign and national enterprises' choice to locate in monotowns, wages are significant and positive, with higher effect on location choice for the foreign firms than for the national ones. Positive effect of wages for all enterprises could imply that higher quality of human capital increases attractiveness of monotowns both for national and foreign business. Concerning the foreign firms, all other variables are insignificant for them. HMP is significant and positive for the national firms. Concerning the national enterprises, the impact

of diversity and urbanization on location choice for the monotowns turned out to be negative; localization effect has an inverted U shape.

As for all other cities, diversity is relatively important for the foreign firms; localization is relatively more important for the national firms. HMP is significant and positive for the foreign firms. Wages are significant and have a negative effect on the location choice both for national and foreign firms. Transport infrastructure and business environment risks are insignificant.

As correlation between diversity and urbanization is high, 78.45% for the sample without Moscow and St. Petersburg, the model is estimated without one of these variables in order to verify the results (Table 5.11 below).

Table 5.11. Enterprise location choice for city: comparing specifications
Dependent variable: enterprise location choice for city

	(1) All tradables-1	(2) All tradables-2	(3) All tradables-3
ldivpq	1.116*** (30.63)		1.116*** (30.63)
lloc	0.243*** (14.37)	0.242*** (14.42)	0.243*** (14.37)
lloc2	-0.00759*** (-12.30)	-0.00265*** (-4.45)	-0.00759*** (-12.30)
lhmecity	0.201*** (6.14)	0.497*** (16.41)	0.201*** (6.14)
lurbpq	0.0285 (0.40)	0.0278 (0.39)	
lwagecity	-0.291*** (-5.76)	-0.802*** (-17.43)	-0.291*** (-5.76)
la_road	-0.132 (-0.39)	-0.128 (-0.38)	-0.0713 (-0.23)
lrw_road	2.449 (1.17)	2.502 (1.19)	2.240 (1.10)
lbusnenvrisk	-0.0492 (-0.61)	-0.0469 (-0.58)	-0.0508 (-0.63)
N	940100	940100	940100
Log lik.	-27925.2	-28711.3	-27925.3

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In Table 5.11 above, Moscow and St. Petersburg were excluded from the specifications, as it was done in the models considered in this chapter. Coefficients remain robust when urbanization variable is excluded. When diversity index is excluded, the significance and signs of the coefficients remain the same. However, the values of optimal localization size, HMP and wage effects increase. Overall, Table 5.11 confirms the choice of the specification used in this chapter, which includes both diversity and urbanization indices.

5.8. Conclusion

This chapter was devoted to the determinants of firms' location choice for a city. The first hypothesis that firms are attracted to the cities with sufficient economic activity of the other firms in various industries is confirmed: diversity measured by the coefficient reflecting both variety and equal distribution of business activity in a city is significant and positive in the majority of specifications. Urbanization level positively affects national firms' choice for a city within agglomerations.

The second hypothesis that firms are attracted to the cities with sufficient economic activity of the other firms in the same industry is confirmed. Localization effects proved to be significant and have an inverted U shape. They are relatively higher for the national firms than for the foreign firms. Optimum is reached at a very low localization level. However, the economies stay positive until localization reaches 80 bn roubles for all firms belonging to the industries producing tradable goods, and around 42 bn roubles for the sample containing only foreign firms. For the monotowns, it is localization level that increases the chances that national firms will work in the town, not diversity or urbanization levels (Table 5.10 above).

The third hypothesis that HMP is significant and positive is confirmed for the majority of samples. Market potential of a city or town is reflected by urbanization and diversity levels; diversity coefficient being positive and significant. Market potential of the other cities is positive and significant too and is much larger for the foreign firms. Concerning city types, for example, for the national firms' location choice for monotown, HMP is significant and positive. Based on the effect of HMP, it can be concluded that foreign firms seem to pursue market seeking strategies.

The fourth hypothesis was that wage effect is ambiguous. Costs are reflected by wages in the chosen specification. Another approach to capture costs, using population density or city area, was tried, but these factors proved to be insignificant. Wage effect is ambiguous as firms have to pay higher wages to better qualified personnel; the same time high quality of human capital contributes to the firms' performance.

National enterprises seem to be attracted by lower wages, while for the foreign enterprises wages turned out to be insignificant in the majority of cases possibly implying interest in qualified workforce. Insignificance of wage for the location decision of the foreign firms leads to the conclusion that they are not likely to pursue efficiency seeking strategies. However, wage effect varies both across types of cities and among national and foreign firms. Both for the national and foreign firms, it positively affects the choice for agglomeration centers and monotowns, and negatively affects the choice for the *other cities*. For the national firms, wage positively affects the choice for a town or city within agglomeration, while for the foreign firms' choice for a town or city within agglomeration wage is insignificant.

The fifth hypothesis was that the level of transport infrastructure development measured with automobile and railroad density positively affects city attractiveness. However, transport infrastructure effect is either insignificant or negative in the majority of cases. Possible reasons are that transport infrastructure, particularly, railroads, are not developed enough, and their location is not efficient enough. The sixth hypothesis, that firms choose cities located in the regions with lower business environment risks, is confirmed for the foreign enterprises. Business environment risks in the region where a city is located are significant and negative for them.

Overall, positive effect of HMP, diversity and localization lead to the conclusion that agglomeration forces seem to prevail over the tendency for more even economic activity distribution. HMP arising from the other cities and regional business risks are more important for the foreign firms than for the national firms; diversity in a city or town is also slightly more important for the foreign firms.

Across industries, for the foreign firms, diversity economies are relatively more important for traditional, basic machinery, and basic materials industries. For the national firms, these economies are also important for science based industries, though the coefficient is lower than for the groups mentioned above. National firms belonging to integrating machinery industry benefit the most from diversity.

For all firms producing tradable goods, localization economies are the most important for science based industries, and lower but significant for the other industries. For the national firms the effects are similar. Foreign firms that benefit from localization economies are those belonging to basic machinery and traditional industries. Localization effects are higher for the foreign firms than for the national firms, if to look at the groups of industries separately. Localization effect keeps the inverted U shape for all groups of industries.

For all firms producing tradable goods, HMP arising from the other cities is the largest for basic machinery industries, followed by basic materials industries, science based and

traditional industries. For the foreign enterprises the effect is the highest for basic machinery industries, followed by traditional and basic materials.

It is confirmed that the foreign enterprises benefit more than the national ones from agglomeration and HMP. It becomes more evident for separate groups of industries. This difference might occur due to more opportunities of location choice that foreign firms have and due to differences in management practices (Vakhitov, 2008).

As for the types of firms, diversity turned out to be relatively more important for the firms with productivity below the mean, while localization proved to be very slightly more important for the firms with productivity above the mean. The difference in HMP is not very large for the two types of firms. For the firms with higher productivity, wage level is insignificant, probably due to the controversial effect of wage discussed above, while lower productivity firms tend to locate in the cities with lower wages.

Less productive firms, firms with state ownership and older firms benefit from diversity more than the other ones. Localization effect is slightly larger for the firms with state ownership and for the 'young' firms. As after 1995 a lot of enterprises were founded based on the assets of the existing enterprises, the difference between the 'young' and 'old' enterprises may not be so explicit. HMP is significant for the firms with private ownership and insignificant for the firms with state ownership. It is significant both for the 'young' and for the 'old' firms, and is larger for the 'young' firms. Private firms are interested in lower wages, while for state firms, wages are insignificant. Both for the 'young' and the 'old' firms, coefficient on wages is significant and negative, being larger for the 'old' firms.

It was assumed that incentives of firms located in agglomeration centers, cities belonging to the agglomeration, monotowns, and the other cities differ. The sign and shape of agglomeration effects are robust in most cases to introducing variables reflecting agglomeration levels for types of cities. Only for monotowns, diversity and urbanization effects become negative. HMP arising from the other cities is either significant and positive, or insignificant; it affects negatively only a choice to locate in an agglomeration center.

Transport infrastructure variables and business environment risks are insignificant for location decisions in the specification for city types. Inclusion of urbanization coefficient did not change the results. However, the coefficient remained persistently insignificant. It might mean that it is not exactly the volume of business activity in the city, but rather its diversity that attracts business or creates favourable conditions for business formation.

Concerning further development of the model, the nested logit model can be estimated, where firstly, choice among the regions will be considered, and then, choice among cities and towns within a region. However, the results are expected to stay similar, as

enterprises are likely to choose mainly among the large cities and because in the estimation presented here regional factors are taken into account.

Conclusion

Agglomeration economies and HMP were found to be important both for the firms' performance and location choice. In Chapters 3 and 4, the impact of agglomeration economies and HMP on the firms' performance was studied, controlling for the size of the firms by including the firms' inputs into the model (labour and fixed assets). In Chapter 4 agglomeration effects were subdivided into those generated by the national and foreign firms.

In Chapter 3, urbanization was found to have a positive impact on the enterprise performance; this effect is robust to changes in specification. Diversity economies are positive as well in all specifications where they are included. Localization economies have an inverted U shape in the majority of specifications. It confirms that congestion is attained at some point of agglomeration process, leading to decrease in agglomeration economies. The share of the same industry enterprises' revenue in a city ('core') is also characterized by an inverted U shape. HMP is significant and positive in the majority of specifications. As it is related to distance and transport costs, improvement of transport infrastructure and decrease in communications costs would probably increase the effects arising from HMP.

Agglomeration effects vary across industries. For the private firms (national and foreign) agglomeration effects are stronger than on average for all tradables, but lower than for the foreign firms. This pattern holds for localization, urbanization and diversity effects, and stays robust to changes in specification. As for different types of cities, concentration of the firms in the same industry (localization) is particularly important for the enterprises in monotowns. The impact of urbanization on enterprise performance is also found to be the highest for the monotowns. HMP arising from the other cities is important, primarily for the towns outside agglomerations, followed by the towns within agglomerations but not agglomeration centers, the monotowns, and the agglomeration centers.

As for the optimal localization level, there are firms located in smaller towns with not so intensive competition and low opportunities for development. These towns are a reasonable focus firstly, for improvement of business climate to attract more firms. Secondly, it seems useful to improve transport infrastructure, particularly with neighbouring territories to increase interaction with them. Firms located close to the optimal localization levels internalize localization economies. There are cities where firms face diminishing localization effects. For them, increasing openness to trade and migration seems to be relevant, in

situation of competition for resources and consumers in a city; it depends on improvement in official regulations, business climate, business services and transport infrastructure.

In Chapter 4, it was found that urbanization economies arising both from national and foreign firms are positive, while those arising from foreign firms are relatively larger in the model with first differencing and GMM. This effect was found both for the sample of national and foreign firms, though for the foreign firms the effects from presence of the other foreign firms in various industries in the city turned out to be larger than for the national firms. The large and robust effects from urbanization can be explained by the increased opportunities for knowledge spillovers, input sharing and to a certain degree by matching (or labour market pooling) in the larger cities, i.e. by the microfoundations behind agglomeration economies.

Inverted U shape is found for localization economies arising from national firms' concentration for the majority of subsamples; the effect of foreign firms' presence in the same industry is contradictory. Positive localization effects start decreasing at extremely low level for certain subsamples. It might be explained by intensive competition on the labour market and low migration levels making job markets limited by cities or agglomerations within commuting distance. However, this result should be treated cautiously, because it is not robust to changes of the subsample (groups of industries, national and foreign firms, etc.). Though diminishing, localization economies stay positive for the actual localization levels in the sample. Theory and empirical evidence on clusters also shows positive effects of same industry firms' co-location.

The impact of the share of the same industry enterprises' revenue in city revenue ('core') is robust for the majority of subsamples and has an inverted U shape. HMP is measured on the city level as HMP arising from the city itself and from the other cities, distances taken into account. HMP arising from the city itself was captured by the urbanization variable. As it was mentioned above, urbanization effect is positive and robust. HMP arising from the other cities is positive in most specifications too. It was compared with HMP on the regional level. On one hand, for the regional level rail road distances were used, making distance measure more realistic. On the other hand, regions in Russia are so large that the regional level HMP is too general to see an effect for a city. HMP arising from the other regions turned out to be positive as well.

In Chapter 5 enterprise location choice for a city was analyzed. The same factors as in Chapters 3 and 4 were considered. The results confirm the assumption that agglomeration level and HMP affect location choice of foreign firms and the probability that national firms will function in a city. Human capital quality was reflected in wages. Their effect is controversial, as it would be expected, because wages reflect both human capital level and

costs that firms bear. However, for the national firms in some industries (traditional and integrating machinery), and for the foreign firms, the impact of wages on firms' performance is positive, i.e. human capital quality outweighs increased costs. Road infrastructure effect is not robust, but for some samples automobile roads effect is positive. The results also confirm the negative effect of regional business environment risks on the firms' performance and the location choice for a city. This outcome was received in most specifications, while in some specifications regional business environment risks were not significant.

In Chapter 5, it was assumed that the foreign firms are more flexible in their location choice for a city than the national firms, and therefore are more demanding for the economic situation and the other factors in the region. It did turn out that foreign firms are more affected by agglomeration economies, HMP and institutional infrastructure. Although nowadays MNCs tend to pursue global strategies combining aspects of market seeking, strategic asset seeking, efficiency seeking, natural resources seeking strategies, it is useful to understand which aspects of MNC strategies prevail in a country.

Estimation shows that foreign enterprises are interested in the large demand, i.e. pursue market seeking strategies. They do not seem to pursue the efficiency seeking strategy as far as wages are concerned. Conclusions about resource seeking strategy cannot be made as the sample contains only manufacturing firms and the relevant variable appeared to be insignificant. Concerning strategic asset seeking, the variable reflecting innovations was insignificant. However, under assumption that there are more innovations in diversified places and places with favourable business climate, the hypothesis that strategic asset seeking strategy may be present cannot be rejected.

In further research, the method of production function estimation proposed by Olley and Pakes (1996) should be considered for the models discussed in chapters 3 and 4, this method being pertinent for the production functions augmented with explicit hypotheses about the determinants of total factor productivity[57].

Besides, in further research, it is planned to measure human capital more precisely as it is assumed to be essential for business performance and location choice. A better measure might be a share of people with higher education in the region. It might be useful to take into account people with vocational training, to consider the quality of workers who do not need higher education to perform their jobs, but whose qualification affects firms' performance. A drawback of the education variables is that only an education degree, not quality of education,

[57] Levinsohn, J., Petrin A. (2003), "Estimating Production Functions Using Inputs To Control For Unobservables," Review of Economic Studies, 2003, 70(2), pp. 317-341;
Olley S. , Pakes A. (1996), "The Dynamics of Productivity in the Telecommunications Equipment Industry", Econometrica, 64(6), pp. 1263-1297.

is taken into account. Besides, it would be useful to construct an index by analogy to human development index reflecting quality of life in cities of different types. A version of such index would be an index of social and economic situation based on the city characteristics available from Rosstat. This index would allow to follow the dynamics of development in various types of cities, and to estimate the effect of city's characteristics on enterprise performance. In this research only spatial dimension of HMP was taken into account. It was assumed that FDI effects quickly attenuate with distance, and therefore omitting the spatial effect will not distort the results significantly. However, it should be checked whether spatial effects of FDI are important (Bode et al., 2009). Another improvement would be to apply a spatial matrix (Anselin, 2005) to reflect more precisely spatial dependence, i.e. spatial interaction among cities.

References

Abel, J.R., Dey I. and T.M. Gabe (2010), Productivity and the density of human capital, *Federal Reserve Bank of New York, Staff Report* No. 440*,* March 2010

Ahrend, R. (2006), Russia's post-crisis growth: its sources and prospects for continuation, *Europe-Asia Studies* 58 (1), 1–24

Ahrend, R. (2008), Understanding Russian Regions' Economic Performance during Periods of Decline and Growth: An Extreme-bound Analysis Approach, *OECD Economics Department Working Papers,* No. 644, OECD Publishing. http://dx.doi.org/10.1787/235866010670

Aitken,B.J., Harrison, A.E.(1999), Do domestic firms benefit from direct foreign investment? Evidence from Venezuela. *American Economic Review*, 89(3), 605-618

Altomonte, C. and L. Resmini (2001), Multinational corporations as catalyst for industrial development. The case of Poland, *LICOS Discussion Paper* No. 9701

Anastassova, L. (2006), Productivity differences and agglomeration across districts in Great Britain. *CERGE-EI Working Paper* No. 289, 34 p.

Andreff, W. (2003), Les multinationales globales. Paris : Editions La Découverte

Andreff, W. (2007), Economie de la transition. La transformation des économies planifiées en économies de marché. Bréal

Anselin, L. (2005), Spatial Econometrics (Ch. 29), *University of Illinois, Urbana-Champaign*, May 20

Audretsch, D. B. (1998), Agglomeration and the location of innovative activity, *CEPR Discussion Paper* No. 1974

Balasubramanyam, V.N., Salisu, M. and Sapsford, D. (1996), Foreign direct investment and growth in EP and IS countries, *Economic Journal, Royal Economic Society*, 106(434), 92-105

Baldwin, R. and T. Okubo (2005), Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting, *NBER Working paper* No. 11650

Barry, F., Holger, G. and E. Strobl (2001), Foreign Direct Investment, Agglomerations and Demonstration Effects, *University College Dublin. Department of Economics, Working paper* No.WP01/04

Barry, F. (2006), Foreign direct investment and the economy of Ireland: some current issues, *University College Dublin,* November

Baum, Ch. (2006), An introduction to modern econometrics using Stata. Texas, A Stata Press Publication

Baum, Ch. F., Schaffer, M. E. and S. Stillman (2007), Enhanced routines for instrumental variables. Generalized method of moments estimation and testing *The Stata Journal,* 7(4), pp. 465-506

Beaudry, C. and A. Schiffauerova (2009), Who's right, Marshall or Jacobs? The localization versus urbanization debate, *Research Policy* 38, 318-337

Békés, G., Kleinert, J. and Toubal, F. (2009), Spillovers from multinationals to heterogeneous domestic firms: evidence from Hungary, *World Economy*, 32, 1408–1433

Belitsky, M. (2011), Driving urban economic growth – evidence from transition economies. *EERC Working paper* No. 11/10E

Benhabib, J. and M. M. Spiegel (1994), The role of human capital in economic development: evidence from aggregate cross-country data, *Journal of Monetary Economics, Elsevier*, 34(2), 143-173

Bernard A., Eaton J., Jensen J., and Kortum S. (2003), Plants and Productivity in International Trade // *American Economic Review,* 93, 1268-1290

Bernard, A.B., Jensen, J.B. and P.K. Schott (2003a), Falling Trade Costs, Heterogeneous Firms, and Industry Dynamics, *Mimeo*

Blomstrom M., Kokko A. (2003), The Economics of Foreign Direct Investment Incentives. *NBER Working Paper* No. 9489

Bode, E., Nunnenkamp, P. and A. Waldkirch (2009), Spatial Effects of Foreign Direct Investment in US States. *Kiel Working Paper* No. 1535

Borensztein, E., J. DeGregorio and J.-W. Lee (1998), How does foreign direct investment affect economic growth? *Journal of International Economics*, 45 (1), 115-35

Boudier-Bensebaa F. (2005), Agglomeration economies and location choice. Foreign direct investment in Hungary, *Economics of Transition,* 13(4), 605–628

Bradshaw, M.J. (2002), The changing geography of foreign investment in the Russian federation. *Russian Economic Trends,* 11 (1), 33–41

Cameron, A.C., Trivedi P.K. (2005), Microeconometrics: methods and applications. Cambridge University Press

Cameron, A.C., Trivedi P.K. (2009) Microeconometrics using STATA. Stata Press. USA

Castellani, D. and F. Pieri (2010), The effect of foreign investments on European regional productivity, *Quaderni del Dipartimento di Economia, Finanza e Statistica* No. 74/2010, Università di Perugia, Dipartimento Economia, Finanza e Statistica

Castiglione, C., Y. Gorbunova, D. Infante, and J. Smirnova (2012), FDI determinants in an idiosyncratic country. A reappraisal over the Russian regions during transition years, *Communist and Post-Communist Studies*

Cheviakova E. Yu. and O.A. Rytchkov (2004), Regional reallocation of Russian industry in transition: a New Economic Geography perspective, *Moscow, Economic Education and Research Consortium, Working Papers Series* No. 04/10, 48 p.

Ciccone, A. and R. E. Hall (1996), Productivity and the density of economic activity, *American Economic Review,* 86(1), 54-70

Combes, P.-P., Lafourcade M., Thisse, J-F. and J-C. Toutain (2008), The rise and fall of spatial inequalities in France: a long-run perspective, *CEPR Discussion Paper* No. 7017

Combes, P.-P., Mayer, T. and J.-F. Thisse (2008a), Economic geography. The integration of regions and nations, *Princeton University Press, Princeton and Oxford*

Crozet, M., Mayer, T., and J.-L. Mucchielli (2004), How do firms agglomerate? A study of FDI in France, *Regional Science and Urban Economics* 34, 27– 54

Damijan, J. P., M. Knell, B. Majcen and M. Rojec (2003), Technology Transfer through FDI in Top-10 Transition Countries: How Important are Direct Effects, Horizontal and Vertical Spillovers? *William Davidson Working Paper* No. 549

Damijan, J. P., M. Knell, B. Majcen and M. Rojec (2013), Impact of firm heterogeneity on direct and spillover effects of FDI: micro evidence from ten transition countries. *Journal of Comparative Economics,* 41(3), 895-922

De Groot, H., Poot, J., and M.J. Smit (2009), Agglomeration externalities, innovation and regional growth: theoretical perspectives and meta-analysis. In: *Handbook of Regional Growth and Development Theories*, *Capello R., and Nijkamp P. (eds.). Cheltenham: Edward Elgar*

De Siano, R. and M. D'Uva (2010), Specialization and growth in Italy: what spatial econometric analysis tells us, *Discussion Paper* No. 1_2010, D.E.S. (Department of Economic Studies), University of Naples "Parthenope", Italy

Dickson, M. (2009), The causal effect of education on wages revisited, *IZA DP* No. 4419

Dolgopyatova T. G. (2008), Ownership concentration and Russian company development: empirical evidence. *Problems of Economic Transition*. 50(5), 7-24

Drapkin, I.M., Kadochnikov, S.M., Fedyunina, A.A. and N.B. Davidson (2012), The impact of FDI on economic diversification: modeling vertical FDI spillovers under firm heterogeneity, *Vestnik UrFU (5),* (In Russian)

Drapkin, I.M., Kadochnikov, S.M., Fedyunina, A.A. and N.B. Davidson (2011a), Efficiency of national companies and diversification of regional industry as factors of external effects from foreign direct investment in the Russian economy, *Russian Journal of Management,* 9(2), 2011, 3-26, (In Russian)

Drapkin, I.M., Davidson N.B., Lopatina T.A. and A.A. Fedyunina (2011b), Factors affecting the activities of foreign companies in Russia, *Journal of Economic Theory*, 2, 85-94, (In Russian)

Drapkin, I.M. (2010), International trade, foreign direct investment and agglomeration effects under heterogeneity conditions. *Izvestiya Ural'skogo Gosudarstvennogo Universiteta*, Seria 3, Social Sciences, (In Russian)

Dunning, J.H. (1993), Multinational enterprises and the global economy, Wokingham: Addison-Wesley

Duranton, G. and D. Puga (2004), Micro-foundations of urban agglomeration economies. In *V. Henderson and J. Thisse, eds. Handbook of Urban and Regional Economics,* 4

Duranton, G. and D. Puga (2001), Nursery cities, *American Economic Review,* 91, 1454-1477

Dyker, D.A. (2004), Caching up and falling behind: post-communist transformation in historic perspective, World Scientific and Imperial College Press

Ellison, G. and E. L. Glaeser (1997), Geographic concentration in U.S. manufacturing industries: a dartboard approach, *The Journal of Political Economy*, 105(5), 889-927

Elmas, F. and S. Degirmen (2009), Turkish Manufacturing Sector during the Financial Crisis, *Anadolu International Conference in Economics*, June 17-19, Eskişehir, Turkey

Elmas, F. and S. Degirmen (2009a), Foreign Direct Investment and Industrial Concentration in the Turkish Manufacturing System, *International Research Journal of Finance and Economics*, 23

Enterprises and markets in 2005-2009: the results of two rounds of survey of the Russian manufacturing enterprises, *Report of State University – Higher School of Economics,* Moscow, 2010, (In Russian) http://www.hse.ru/news/hse_pubs/17619989.html

Eriksson, R. (2010), Localized spillovers and knowledge flows: how does proximity influence the performance of plants? *Papers in Evolutionary Economic Geography (PEEG)* No 1004, Utrecht University, Section of Economic Geography, revised March 2010

Expert (2008), Ranking of regional investment attractiveness. Analytical agency 'Expert' http://www.raexpert.ru/ratings/regions/ratingclass/. Concept of the project: http://www.raexpert.ru/ratings/regions/concept/

Foundations of transition economy theory (1996). Moscow State Institute of International Relations, 320 p. (In Russian)

Fujita, M. (1988), A monopolistic competition model of spatial agglomeration: a differentiated product approach, *Regional Science and Urban Economics* 18, 87-124

Glaeser, E. L. and J. D. Gottlieb (2009), The wealth of cities: agglomeration economies and spatial equilibrium in the United States. *NBER Working paper* No.14806

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A. and A. Shleifer (1992), Growth in cities, *The Journal of Political Economy,* 100(6), Centennial Issue (Dec., 1992), 1126-1152

Golovanova, S.V. (2008), Endogenous change of concentration and territorial allocation of production in Russia, *Working paper* No. 1/2008/01, Institutional problems of the Russian economy, Moscow, SU-HSE, (In Russian)

Gonchar, K. and Ph. Marek (2013), Natural-resource or market-seeking FDI in Russia? An empirical study of locational factors affecting the regional distribution of FDI entries, *Discussion paper* No. 3, Halle Institute for Economic Research

Gonchar, K. (2010), Urban agglomeration effects and company productivity in Russia: Empirical Evidence Based on Manufacturing Industry Survey, *Report at ISNIE*

Gonchar, K. (2008), Agglomeration effects and industrial competitiveness, Series 'Conferences of the Department of Economics, Ural State University', 5th International Conference, Ekaterinburg, April

Granberg, A.G. (2001), Foundations of regional economics, 2nd edition, Moscow, State University-Higher School of Economics – 495 p. (In Russian)

Griffith R., Simpson H. (2003), Characteristics of foreign-owned firms in British manufacturing. *NBER Working Paper* No. 9573

Hall, R.E. and Ch. I. Jones (1999), Why do some countries produce so much more output per worker than others? *The Quaterly Journal of Economics*, 114(10), 83-116

Hanousek, J., Kocenda, E. and M. Masika (2012), Firm efficiency: domestic owners, coalitions, and FDI, *CERGE-EI Working paper* No. 456, Prague, April

Harris, C. (1954), The Market as a factor in the localization of industry in the United States, *Annals of the Association of American Geographers,* 64, 315–348

Head, K. and Th. Mayer (2004), Market potential and the location of Japanese investment in the European Union, *The Review of Economics and Statistics,* 86(4), 959-972

Helpman, E. (2005), Trade, FDI, and the Organization of Firms, Harvard University and CIAR, September 26

Helpman E., Melitz, M. and Yeaple, S. (2004), Export versus FDI with Heterogeneous Firms, *American Economic Review*, 94, 300-316

Helpman, E., Melitz, M. and Y. Rubinstein (2007), Estimating Trade Flows: Trading Partners And Trading Volumes, *Quaterly Journal of Economics*, MIT Press, 123(2), 441-487

Henisz, W. J. (2009), Concentrated Power, Foreign Direct Investment and Economic Growth, 13th Annual Conference of The International Society for New Institutional Economics (ISNIE), Conference paper

Henderson, J.V., A. Kuncoro and M. Turner (1995), Industrial development in cities, *Journal of Political Economy* 103, 1067-1085

Henderson, J.V. (2003), Marshall's scale economies, *Journal of Urban Economics,* 53, 1-28

Jacobs, J. (1961), The death and life of great American cities, New York, Random House

Jacobs, J. (1969), The economies of cities, New York, Random House

Kadochnikov, S. M and P.V. Esin (2006), The factors of product innovations in restructuring of the Russian companies (based on the example of the Urals region), *Russian Journal of Management.* 4(1), 29-54, (In Russian)

Kadochnikov, S. M and T. V. Kulakova (2004), Foreign direct investment in modern theory of institutional economics and in international trade theory: the main approaches and empirical analysis. Textbook. Ekaterinburg, Publishing house of the Urals University, (In Russian)

Kadochnikov, S. and I. M. Drapkin (2007) The effects and factors of foreign direct investment impact on the host economy, *Journal of Economic Theory*, 3, 31-45, (In Russian)

Kadochnikov, S., Vorobev, P. Legkaya, Y. and N. Davidson (2008) Effects of Regional FDI Concentration on Enterprise Performance in Russia. Report at the Conference, State University – Higher School of Economics, Moscow, April

Kathuria, V. (2000), Productivity spillovers from technology transfer to Indian manufacturing firms. *Journal of International Development*, 12, 343–369

Kireev, A.P. (2011), Theories of international trade, XXI century. *Economic school. Almanac*, 7, International Economics, Editors: Kireev A.P., Matveenko, V.D., St. Petersburg, Economic School, SU-HSE

Kolomak, E.A. (2006), Managing competitive environment at the sub-federal level in Russia, *Russian Academy of Sciences, Siberian Brunch, Institute for Economics and Organization of Manufacturing Production,* Novosibirsk

Krugman, P. (1980), Scale economies, product differentiation, and the pattern of trade, *American Economic Review,* 70, 950-9

Krugman, P. (1991), Increasing returns and economic geography, *Journal of Political Economy*, 99, 483-499

Krugman, P. (1998), The role of geography in development, *Annual World Bank Conference on Development Economics*, April

Lappo, G. and P. Polyan (2007), Results of urbanization in Russia by the end of XX century. Center for Strategic Development North-West, http://www.ecsman.edu.ru/Economics, Sociology, Management, (In Russian)

Ledyaeva, S. (2009), Spatial econometric analysis of foreign direct investment determinants in Russian regions, *The World Economy,* 32(4), 643–666

Ledyaeva, S, Karhunen, P. and R. Kosonen (2010), Location choice of Chinese, Japanese and the US FDI in Russia: Do cross-country agglomeration effects matter? For Nordic Conference in Development Economics

Lin, P. and K. Saggi (2007), "Multinational firms, exclusivity, and the degree of backward linkages", *Journal of International Economics*, 71, 206 - 220

Lipsey, R.E. (2002), Home and host country effects of FDI, *NBER Working Paper* No. 9293

Mariev, O. S. and D. V. Nesterova (2011), Factors of attracting FDI into the Russian regions, *Ekonomika regiona*, 4

Marinova S.T., Marinov M.A. (2003), Foreign direct investment in Central and Eastern Europe, Ashgate

Markusen. J., and Venables, A.J. (1999), Foreign direct investment as a catalyst for industrial development, *European Economic Review*, 43, 335-356.

Markusen, J. and N. Trofimenko (2007), Teaching locals new tricks: foreign experts as a channel of knowledge transfers, *NBER Working paper* No. 12872

Marshall, A. (1920), Principles of Economics, 8th edition, London, Macmillan and Co. Original edition 1890

Martin, P. (1998), Can Regional Policies Affect Growth and Geography in Europe, *World Economy,* 21(6), 757-774

Martin, P., Mayer, T. and F. Mayneris (2011), Spatial concentration and plant-level productivity in France, *Journal of Urban Economics*, 69, 182–195

Mayer, Th. Mejean, I. and B. Nefussi (2010), The location of domestic and foreign production affiliates by French multinational firms, *IMF Working Paper* No. WP/10/22

McFadden, D. (1974), Conditional logit analysis of qualitative choice behavior, Chapter 4 In P. Zarembka (ed.), Frontiers of Econometrics, New York, Academic Press

McFadden, D. (1978), Modelling the Choice of Residential Location, in Karlquist, A. et al. (eds), *Spatial Interaction Theory and Residential Location,* Amsterdam: North-Holland, 75-96

Melitz, M. (2003), The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity, *Econometrica* 71, 1695-1725

Melitz, M. and Ottaviano G. (2008), Market Size, Trade, and Productivity, *Review of Economic Studies,* 75, 295–316

Merlevede B. and K. Schoors (2008), How and by how much does foreign direct investment increase the productivity of domestic firms? *Mimeo*

Mills, E. S. (1967), An aggregative model of resource allocation in a metropolitan area, *American Economic Review Papers and Proceedings,* 57(2), 197–210

Mikhailova, A.A. (2009), The role of FDI in the economy of the countries-recipients, *Ekonomicheskaya nauka sovremennoi Rossii*, 3(46), 84-93, (In Russian)

Mikhailova, T. (2011), Looking for multiple equilibiria in Russian urban system, *EERC Working paper* No. 10/08E

Mikhailova, T. (2013), Looking for multiple equilibria in Russian urban system. Presentation at the 2nd International Conference 'Industrial Organization and Spatial Economics', St. Petersburg, Center for Market Studies and Spatial Economics National Research University 'Higher School of Economics', 27-30 June

Mirrlees, J. A. (1972), The optimum city, *Swedish Journal of Economics,* 74(1), 114–135

Monotowns in Russia: how to survive crisis? Analysis of social and economic problems of monotowns in the context of the world financial and economic crisis, affecting the situation in the city-forming corporations, *Institute for regional policy,* Moscow, 2008, (In Russian)

Navaretti, G.B. and A.J. Venables (2004), Multinational firms in the world economy, Princeton and Oxford: *Princeton University Press*

Neffke, F., Henning, S. M., Boschma, R., Lundquist, K.-J. and L.-O. Olander (2008), Who needs agglomeration? Varying agglomeration externalities and the industry life cycle, *Working Paper,* March 2008. Utrecht University, the Netherlands

Neffke, F. (2009), Productive places. The influence of technological change and relatedness on agglomeration externalities, Faculty of Geosciences, Utrecht University, the Netherlands

Okubo, T. (2006), Trade Liberalisation and Agglomeration with Firm Heterogeneity – Forward and Backward Linkages, Mimeo

Okubo, T. and E. Tomiura (2010), Productivity distribution, firm heterogeneity, and agglomeration: evidence from firm-level data, *Discussion paper* No 10017, Research Institute of Economy, Trade and Industry (RIETI)

Ottaviano, G. and J. Thisse (2004), Agglomeration and Economic Geography, in *V. Henderson and J. Thisse, eds. Handbook of Urban and Regional Economics,* 4

Pavitt, K. (1984), Sectoral patterns of technical change: towards a taxonomy and a theory, *Research Policy*, 13, 343-373

Population census, Rosstat, 2010

http://www.gks.ru/free_doc/new_site/perepis2010/croc/perepis_itogi1612.htm

Porter, M.E. (2000), Location, competition, and economic development: local clusters in a global economy, *Economic Development Quarterly* 14 (1), 15-34

Quigley (1998), Urban Diversity and Economic Growth, *Journal of Economic Perspectives*, 12(2) Spring, 127-138

Robert-Nicoud, F. (2013), Agglomeration economies: matching and sharing, *Presentation at CMSSE Summer School in Spatial Economics and Imperfect Markets: empirical aspects. Pushkin, 3-16 July*

Rodriguez-Clare, A. (1996), Multinationals, linkages, and economic development, *American Economic Review*, 86, 852 – 873

Rosenthal, S. S. and W. C. Strange (2003), Geography, industrial organization and agglomeration, *Review of Economics and Statistics,* 85(2), 377-393

Rosenthal, S. S. and W. C. Strange (2004), Evidence on the nature and sources of agglomeration economies, *in V. Henderson and J. Thisse, eds. Handbook of Urban and Regional Economics,* 4

Russian Economic Report, The World Bank in Russia: Economic Review, June 2007

Rutherford, Th., Tarr D. G. and O. Shepotylo (2005), The Impact on Russia of WTO Accession and The Doha Agenda: the importance of liberalization of barriers against foreign direct investment in services for growth and poverty reduction," in Thomas Hertel and L. Alan Winters (eds.), *Putting Development Back into the Doha Agenda: Poverty Impacts of a WTO Agreement*

Rutherford, T., Tarr, D. (2008), Poverty effects of Russia's WTO accession: modeling «real» households with endogenous productivity effects, *The Journal of International Economics,* 75, 131-150

Rutkowski, A. (2006), Inward FDI, concentration, and profitability in the CEECs: were the domestic firms crowded out or strengthened? *Transnational Corporations,* 15(3)

Sabirianova, K., J. Svejnar and K. Terrell (2005), Foreign Investment, corporate ownership, and development: are firms in emerging markets catching up to the world standard?, *William Davidson Institute Working Paper* No. 734*, William Davidson Institute at the University of Michigan.*

Schaffer, M.E. (2010) xtivreg2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models. http://ideas.repec.org/c/boc/bocode/s456501.html

Shastitko A., Plaksin S. and E. Yakovleva (2007), Investment image of Russia, *Strategies of Russia*, 4, 73-83

Shiltsin E.A. (2010), Estimation of convergence and divergence of regional space in Russia and Siberia. *Dissertation for the degree of Candidate of Sciences in Economics, Novosibirsk, Sibirian branch of the Russian Academy of Sciences, Institute of Economics and Organization of Manufacturing Production,* (In Russian)

Shepotilo, O. (2011), Cities in transition. *EERC Working paper* No. 11/08E

Smarzynska-Javorcik, B. (2004), Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages, *American Economic Review*, 94, 605 - 627

Stancik J. (2007), Horizontal and Vertical FDI Spillovers: Recent Evidence from the Czech Republic, *CERGE-EI Working paper* No. 340

Sveikauskas L. (1975), The Productivity of Cities, *The Quarterly Journal of Economics*, 89(3), 393-413

Vakhitov V. (2008), Agglomeration economies and geographic concentration of manufacturing in Ukraine, Dissertation, Director: Dr. Christopher Bollinger, Professor of Economics, Lexington, Kentucky

Vakhitov V. and C. Bollinger (2010), Effects of ownership on agglomeration economies: evidence from Ukraine, *Discussion Paper* No. 24, Kiev School of Economics

Venables, A.J. (1996), Equilibrium locations of vertically linked industries, *International Economic Review,* 37, 341-359

Vernon, R. (1960), Metropolis 1985, Cambridge, MA, Harvard University Press

Vitali, S., Napoletano, M. and G. Fagiolo (2009), Spatial localization in manufacturing: a cross-country analysis. LEM (Laboratory of Economics and Management), *Sant' Anna School of Advanced studies, Working paper* No. 4

Vorobyev P.V., Kislyak N.V., Davidson N.B. (2010), Spatial concentration and firm performance in Russia, Kiev, EERC, 52 p.

Vorobyev P.V., Aleynikova I.S., Isakidis V.A. (2008), Models of regional industrial clusters' organization: the review of international experience, Ekaterinburg, *Center for Regional Economic Studies, Department of Economics, Ural State University,* 31 p. (In Russian)

Vorobyev, P. and A. Scherbinina (2009), Monotowns: the choice of industries for diversification, *Centre for regional economic studies, Department of Economics, Ural State University,* Yekaterinburg, (In Russian)

UNCTAD (2003), World Investment Report 2003 – FDI Policies for Development: National and International Perspectives, *United Nations Publication*

UNCTAD (2011), World Investment Report 2011 - Non-Equity Modes of International Production and Development, *United Nations Publication*

UNCTAD (2012), World Investment Report 2012 - Towards a New Generation of Investment Policies, United Nations Publication

UNCTAD (2013), World Investment Report 2013: Annex Tables http://unctad.org/en/Pages/DIAE/World%20Investment%20Report/Annex-Tables.aspx

Weber, A. (1929), The theory of the location of industries, Chicago University Press, Chicago (English translation from the German version published in 1909)

Wooldridge, J. M. (2002), Econometric analysis of cross section and panel data, MIT Press

WTO (2012), World Trade Organization - Statistics database, URL: http://stat.wto.org/CountryProfile/WSDBCountryPFView.aspx?Language=E&Country=RU

Yeaple, S. (2002), A Simple Model of Firm Heterogeneity, International Trade, and Wages, Mimeo

Yudaeva K., K. Kozlov, N. Melentieva and N. Ponomareva (2003), Does foreign ownership matter? The Russian experience, *Economics of Transition* 11(3), 383-409

Zubarevich, N. (2010), Regions of Russia: inequality, crisis, modernization, Independent Institute for Social Policy, Moscow, (In Russian)

Appendix A. Correlation matrices

Table A.1. Correlation matrix for Chapter 3. Tradable industries. Enterprises, for which *city* level data is present. 2002-2008.

	lrevenue	lfixed_assets	llabour	core	core2	lloc	lloc2	ldivpq	lurbpq	lhmecity	lwagecity	la_road	lrw_road	lbusnenvrisk
lrevenue	1.0000													
lfixed_assets	0.7830	1.0000												
llabour	0.8296	0.7671	1.0000											
core	0.2755	0.2377	0.2289	1.0000										
core2	0.2016	0.1832	0.1638	0.9145	1.0000									
lloc	0.0977	0.0267	0.0196	0.1995	0.0896	1.0000								
lloc2	0.1164	0.0392	0.0179	0.2418	0.1052	0.9727	1.0000							
ldivpq	0.0462	-0.0169	-0.0418	-0.2679	-0.3041	0.5398	0.5978	1.0000						
lurbpq	0.1186	0.0261	-0.0245	-0.1975	-0.2250	0.5430	0.6248	0.8950	1.0000					
lhmecity	0.0915	0.0222	0.0139	-0.0004	0.0051	-0.0918	-0.1104	-0.1856	-0.0783	1.0000				
lwagecity	0.1865	0.0765	-0.0280	0.0493	0.0301	0.2621	0.3276	0.2712	0.5340	0.4148	1.0000			
la_road	-0.0073	-0.0453	-0.0533	-0.0325	-0.0939	0.2184	0.2734	0.3919	0.3697	0.1746	0.0866	1.0000		
lrw_road	0.0269	-0.0324	-0.0382	-0.0223	-0.0887	0.2884	0.3557	0.4800	0.4812	0.1278	0.1230	0.8830	1.0000	
lbusnenvrisk	-0.0708	-0.0088	0.0155	-0.0034	0.0363	-0.2882	-0.3420	-0.4175	-0.4682	-0.0244	-0.1886	-0.4149	-0.4685	1.0000

Table A.2. Correlation matrix for Chapter 3. Tradable industries. Enterprises, for which *region* level data is present. 1999-2008.

	lrevenue	lfixed_assets	llabour	core	core2	lloc	lloc2	ldivpq	lgrp	lhmeR	lwage	la_road	lrw_road	lbusnenvrisk
lrevenue	1.0000													
lfixed_assets	0.7902	1.0000												
llabour	0.8313	0.7876	1.0000											
core	0.2796	0.2617	0.2624	1.0000										
core2	0.2243	0.2166	0.2130	0.9458	1.0000									
lloc	0.1927	0.0989	0.1035	-0.0331	-0.0495	1.0000								
lloc2	0.2035	0.1047	0.0972	-0.0193	-0.0492	0.9792	1.0000							
ldivpq	0.1726	0.0858	0.0955	-0.4402	-0.3947	0.6663	0.6844	1.0000						
lgrp	0.2516	0.1261	0.0505	-0.0721	-0.0731	0.4459	0.5356	0.4806	1.0000					
lhmeR	0.1986	0.0788	0.0344	-0.0198	-0.0210	0.0289	0.0489	0.0216	0.2922	1.0000				
lwage	0.2767	0.1226	0.0105	-0.0144	-0.0153	0.2296	0.2849	0.1799	0.7126	0.6365	1.0000			
la_road	0.0108	-0.0263	-0.0047	-0.0937	-0.1080	0.2091	0.2559	0.3201	0.2529	0.3667	-0.0162	1.0000		
lrw_road	0.0239	-0.0244	0.0016	-0.0957	-0.1092	0.2530	0.3061	0.3601	0.3111	0.2824	-0.0005	0.8815	1.0000	
lbusnenvrisk	-0.0616	-0.0142	-0.0122	0.0887	0.0950	-0.3247	-0.3705	-0.4047	-0.4275	-0.0056	-0.0556	-0.4598	-0.4704	1.0000

Table A.3. Correlation matrix of the variables employed in the empirical model - for Chapter 4. Tradable industries. Enterprises, for which *city* level data is present. 2002-2008.

	lrevenue	lfixed_assets	llabour	lhmecity	lwagecity	core	core2	llocnl	llocnl2	lurbnl	llocf	lurbf	a_road	rw_road	lbusnenvrisk
lrevenue	1.0000														
lfixed_assets	0.7897	1.0000													
llabour	0.8394	0.7724	1.0000												
lhmecity	0.1218	0.0474	0.0469	1.0000											
lwagecity	0.1884	0.0767	-0.0136	0.4534	1.0000										
core	0.2840	0.2464	0.2408	0.0081	0.0422	1.0000									
core2	0.2041	0.1853	0.1702	0.0074	0.0215	0.9124	1.0000								
llocnl	0.0002	-0.0547	-0.0628	-0.1046	0.2364	0.1562	0.0685	1.0000							
llocnl2	0.0261	-0.0361	-0.0556	-0.1151	0.2989	0.2011	0.0853	0.9756	1.0000						
lurbnl	-0.0454	-0.0971	-0.0921	-0.0674	0.2758	-0.1317	-0.1284	0.4449	0.4610	1.0000					
llocf	0.0788	0.0112	-0.0386	-0.1915	0.3445	0.1073	0.0117	0.4923	0.5929	0.3756	1.0000				
lurbf	0.1068	0.0387	-0.0196	-0.1748	0.3396	0.0501	-0.0293	0.4580	0.5497	0.2358	0.9246	1.0000			
a_road	0.0439	-0.0118	-0.0631	0.1025	0.3960	0.0156	-0.0564	0.2772	0.3545	0.2682	0.5299	0.5207	1.0000		
rw_road	0.0492	-0.0183	-0.0618	-0.0063	0.3097	0.0125	-0.0671	0.3469	0.4331	0.3151	0.6530	0.6515	0.8727	1.0000	
lbusnenvrisk	-0.0466	0.0008	0.0244	-0.0104	-0.0969	0.0006	0.0326	-0.2434	-0.2830	-0.2023	-0.4074	-0.4166	-0.3113	-0.4406	1.0000

Table A.4. Correlation matrix of the variables employed in the empirical model with productivity - for Chapter 4.

	productivity	lfixed_assets	llabour	lhmecity	lwagecity	core	core2	llocnl	llocnl2	lurbnl	llocf	lurbf	a_road	rw_road	lbusnenvrisk
productivity	1.0000														
lfixed_assets	0.0341	1.0000													
llabour	-0.0440	0.7724	1.0000												
lhmecity	0.0022	0.0474	0.0469	1.0000											
lwagecity	0.0307	0.0767	-0.0136	0.4534	1.0000										
core	0.0953	0.2464	0.2408	0.0081	0.0422	1.0000									
core2	0.1184	0.1853	0.1702	0.0074	0.0215	0.9124	1.0000								
llocnl	0.0108	-0.0547	-0.0628	-0.1046	0.2364	0.1562	0.0685	1.0000							
llocnl2	0.0174	-0.0361	-0.0556	-0.1151	0.2989	0.2011	0.0853	0.9756	1.0000						
lurbnl	-0.0043	-0.0971	-0.0921	-0.0674	0.2758	-0.1317	-0.1284	0.4449	0.4610	1.0000					
llocf	0.0286	0.0112	-0.0386	-0.1915	0.3445	0.1073	0.0117	0.4923	0.5929	0.3756	1.0000				
lurbf	0.0182	0.0387	-0.0196	-0.1748	0.3396	0.0501	-0.0293	0.4580	0.5497	0.2358	0.9246	1.0000			
a_road	-0.0105	-0.0118	-0.0631	0.1025	0.3960	0.0156	-0.0564	0.2772	0.3545	0.2682	0.5299	0.5207	1.0000		
rw_road	-0.0111	-0.0183	-0.0618	-0.0063	0.3097	0.0125	-0.0671	0.3469	0.4331	0.3151	0.6530	0.6515	0.8727	1.0000	
lbusnenvrisk	0.0008	0.0008	0.0244	-0.0104	-0.0969	0.0006	0.0326	-0.2434	-0.2830	-0.2023	-0.4074	-0.4166	-0.3113	-0.4406	1.0000

Table A.5. Correlation matrix for Chapter 5. Tradable industries. 2007.

	ldivpq	core	core2	lloc	lloc2	lhmecity	lurbpq	lwagecity	la_road	lrw_road	lbusnenvrisk
ldivpq	1.0000										
core	-0.2822	1.0000									
core2	-0.3172	0.9187	1.0000								
lloc	0.5275	0.2051	0.0937	1.0000							
lloc2	0.5873	0.2428	0.1076	0.9731	1.0000						
lhmecity	-0.2443	-0.0057	-0.0012	-0.1563	-0.2023	1.0000					
lurbpq	0.9082	-0.2155	-0.2436	0.5338	0.6147	-0.2725	1.0000				
lwagecity	0.5075	0.0606	0.0212	0.4119	0.4951	-0.3009	0.6906	1.0000			
la_road	0.4119	-0.0396	-0.1017	0.2292	0.2849	0.1573	0.4044	0.1810	1.0000		
lrw_road	0.4854	-0.0402	-0.1037	0.2940	0.3579	0.1846	0.5056	0.2956	0.8874	1.0000	
lbusnenvrisk	-0.4154	0.0122	0.0454	-0.2866	-0.3376	-0.0142	-0.4765	-0.3580	-0.3596	-0.4660	1.0000

Table A.6. Correlation matrix for Chapter 5. Tradable industries. Sample without Moscow and St. Petersburg. 2007.

	ldivpq	core	core2	lloc	lloc2	lhmecity	lurbpq	lwagecity	la_road	lrw_road	lbusnenvrisk
ldivpq	1.0000										
core	-0.4093	1.0000									
core2	-0.3801	0.9375	1.0000								
lloc	0.2799	0.2305	0.1572	1.0000							
lloc2	0.2757	0.2840	0.2040	0.9843	1.0000						
lhmecity	0.0433	-0.0045	-0.0315	0.0268	0.0280	1.0000					
lurbpq	0.7845	-0.3769	-0.3216	0.2579	0.2672	0.0851	1.0000				
lwagecity	-0.0305	0.1025	0.1137	0.1112	0.1337	-0.0254	0.2507	1.0000			
la_road	0.0456	-0.0525	-0.0779	-0.0361	-0.0529	0.5643	-0.1027	-0.3977	1.0000		
lrw_road	0.0384	-0.0396	-0.0661	-0.0230	-0.0353	0.6090	-0.0618	-0.2828	0.8670	1.0000	
lbusnenvrisk	0.0330	-0.0394	-0.0239	0.0149	0.0141	-0.1907	-0.0109	-0.0585	-0.4360	-0.3002	1.0000

Appendix B. Classification into five groups of industries[58]

1 - traditional goods

Manufacture of tobacco products, Finishing of textiles, Manufacture of made-up textile, articles, except apparel, Manufacture of other textiles, Manufacture of knitted and crocheted articles, Manufacture of leather clothes; Manufacture of other wearing apparel and accessories; Dressing and dyeing of fur; manufacture of articles of fur; Tanning and dressing of leather; Manufacture of luggage, handbags and the like, saddlery and harness; Manufacture of footwear; Sawmilling and planing of wood; impregnation of wood; Manufacture of veneer sheets; manufacture of plywood, laminboard, particle board, fibre board and other panels and boards; Manufacture of builders carpentry and joinery; Manufacture of wooden containers; Manufacture of other products of wood; manufacture of articles of cork, straw and plaiting materials; Printing and service activities related to printing; Reproduction of recorded media; Cutting, shaping and finishing of stone; Manufacture of furniture; Manufacture of jewellery and related articles; Manufacture of musical instruments; Manufacture of sports goods; Manufacture of games and toys; Miscellaneous manufacturing n.e.c.

2 - basic materials

Preparation and spinning of textile fibres; Textile weaving; Manufacture of knitted and crocheted fabrics; Manufacture of pulp, paper and paperboard; Manufacture of articles of paper and paperboard; Publishing; Manufacture of coke oven products; Manufacture of refined petroleum products; Manufacture of basic chemicals; Manufacture of paints, varnishes and similar coatings, printing ink and mastics; Manufacture of pharmaceuticals; Manufacture of soap and detergents, cleaning and polishing preparations, perfumes and toilet preparations; Manufacture of other chemical products; Manufacture of man-made fibres; Manufacture of rubber products; Manufacture of plastic products; Manufacture of glass and glass products; Manufacture of ceramic goods other than for construction purposes; Manufacture of ceramic tiles and flags; Manufacture of bricks, tiles and construction products, in baked clay; Manufacture of cement, lime and plaster; Manufacture of articles of concrete, plaster and cement; Manufacture of other non-metallic mineral products; Manufacture of cast iron and steel and of ferro-alloys, of rolled metal; Manufacture of cast iron and steel tubes; Other first processing of iron and steel n.e.c.; Manufacture of non-ferrous metals; Casting of metals; Manufacture of structural metal products; Manufacture of tanks, reservoirs and containers of metal; manufacture of central heating radiators and boilers; Manufacture of steam generators, except central heating hot water boilers, of nuclear reactors; Forging, pressing, stamping and roll forming of metal; powder metallurgy; Treatment and coating of metals; general mechanical engineering; Manufacture of cutlery, tools and general hardware; Manufacture of other fabricated metal products; Recycling of metal waste and scrap; Recycling of non-metal waste and scrap; Manufacture of gas; distribution of gaseous fuels through mains.

3 - basic machinery

Manufacture of machinery for the production and use of mechanical power; Manufacture of other general purpose machinery; Manufacture of agricultural and forestry machinery; Manufacture of machine tools; Manufacture of other special purpose machinery; Manufacture of domestic appliances n.e.c.; Manufacture of electric motors, generators and transformers;

[58] Only tradable goods are included: division D (manufacturing), partly E (electricity) and partly K (real estate, renting and business activities).
Divisions A (agriculture, hunting and forestry), B (fishing), C (mining and quarrying) of tradable goods are not included, because we assume that agglomeration economies are not so important for them.
Based on 3-digit industrial classification (OKVED).

Manufacture of electricity distribution and control apparatus; Manufacture of insulated wire and cable; Manufacture of accumulators, primary cells and primary batteries; Manufacture of lighting equipment and electric lamps; Manufacture of electrical equipment n.e.c.; Manufacture of television and radio transmitters and apparatus for line telephony and line telegraphy; Manufacture of television and radio receivers, sound or video recording or reproducing apparatus and associated goods; Manufacture of bodies (coachwork) for motor vehicles; manufacture of trailers and semi-trailers; Manufacture of parts and accessories for motor vehicles and their engines.

4 – integrating machinery

Manufacture of motor vehicles; Building and repairing of ships and boats; Manufacture of railway and tramway locomotives and rolling stock; Manufacture of motorcycles and bicycles; Manufacture of other transport equipment n.e.c.

5 – science based

Processing of nuclear fuel; Manufacture of weapons and ammunition; Manufacture of office machinery and computers; Manufacture of electronic valves and tubes and other electronic components; Manufacture of medical and surgical equipment and orthopedic appliances; Manufacture of instruments and appliances for measuring, checking, testing, navigating and other purposes; Manufacture of industrial process control equipment; Manufacture of optical instruments and photographic equipment; Manufacture of watches and clocks; Manufacture of aircraft and spacecraft; Software consultancy and supply; Database activities; Research and experimental development on natural sciences and engineering; Research and experimental development on social sciences and humanities; Architectural and engineering activities and related technical consultancy, finding technical solutions n.e.c.; Technical testing and analysis.

Appendix C. Data Sources

Data	Source
Firm level data	SPARK-Interfax[59] http://www.spark-interfax.ru/Front/Index.aspx
Gross regional product, regional average wage, transport infrastructure	Federal service of state statistics (Rosstat) regional data www.gks.ru
Regional investment climate indices • Regional investment potential[60] • Regional investment risk[61] (used in this research to reflect regional business environment risks)	Analytical agency 'Expert'[62] http://www.raexpert.ru/ratings/regions/ratingclass/ *the results of Ranking of investment attractiveness of the Russian regions are published in the journal 'Expert' annually since 1996*
City social and economic indicators, including city average wage	Rosstat city level data http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/publications/catalog/doc_1138631758656

[59] The data source SPARK (System for professional analysis of markets and companies) is compiled by the Interfax Information Services Group and includes all official data coming from more than ten sources, such as Federal service of state statistics (Rosstat), Federal tax service (FNS), Federal service for financial markets (FSFR), Bank of Russia.

[60] *Regional investment potential* (a quantitative characteristics), contains nine potentials (until 2005 there were eight): natural resources; labour; production; innovation; institutional; infrastructure; financial; consumption; tourist.

[61] *Regional investment risk* (a qualitative characteristics reflecting probability to lose investment and income from investment), contains seven risks: economic, financial, social, ecological, criminal, legislative, governance.

[62] The main data sources for the ranking are the following: data compiled by Rosstat, Ministry of Finance of the Russian Federation, Ministry of economic development and trade of RF, Central Bank of RF, Ministry of taxation and levies of RF, Ministry of natural resources of RF, Center for economic environment (Government of RF), legislation database 'Consultant Plus-Region', database of the ranking agency 'Expert RA'. Besides, information provided by certain subjects of federation (from their official web sites or sent upon request) is used.
Estimation of weighs of each component in the total potential and integral risk is based on the annual surveys among the experts from the Russian and foreign investment, consulting companies and enterprises.