The Valuation of Local Government Spending: Gravity Approach and Aggregate Implications[*]

Wookun Kim[†]

February 18, 2020

[Click Here for the Most Recent Version](#)

Abstract

How much do people value local government spending? What are the effects of fiscal transfers that finance this spending? I develop a spatial equilibrium framework where people's simultaneous (internal) migration and commuting choices reveal preferences. I combine this framework with administrative data from South Korea and leverage plausibly exogenous variation in local government spending across districts induced by national tax reforms in 2008 and 2012. The estimated mobility responses imply that workers value each additional dollar of per-capita local government spending by 75 cents of their after-tax income. The general-equilibrium counterfactuals imply that a fiscal arrangement with lower redistribution would result in aggregate gains. A key aspect of my analysis is that bilateral migration and commuting decisions are made jointly. Ignoring any of these margins biases the estimates of preferences for public goods, and of distance elasticities of migration or commuting which play a central role in quantitative spatial models.

Keywords: gravity, migration, commuting, local public goods, intergovernmental transfers
JEL codes: H3, H77, J61, R12, R13, R5

---

1 Introduction

How much do people value local government goods? Answering this question is important to inform questions in public finance and urban economics. In this paper, I implement a new gravity approach to measure these preferences. I implement this approach in the context of South Korea, a highly decentralized economy with large heterogeneity in local spending across across districts. Then, I compute the optimal levels of fiscal redistribution implied by this measurement.

In the spirit of Tiebout (1956), my approach to estimate the valuation of local government spending is based on mobility. The key novel feature is that preferences for government spending are revealed by people's bilateral migration and commuting decisions in a context where moving is costly. Accounting jointly for both margins of mobility—migration and commuting—is important because these choices are linked and may respond to government spending. Workers may move to places with generous provision of local government goods, but they may also find places attractive to live in if they facilitate access to jobs via commuting. Furthermore, the location of origin (i.e., from where a worker migrates) may influence the choice of both residence and workplace. Using a gravity equation that captures these margins and quasi-natural variation in government spending, I find that the marginal valuation of a dollar of local government spending is equal to 75 cents of disposable income. A key takeaway from my analysis is that ignoring any of these dimensions (place of origin, place of residence, and workplace) biases the estimates of preferences for public goods and of distance elasticities of migration or commuting which play a central role in quantitative spatial models.

The empirical setting of this paper is South Korea. There are three key aspects of the South Korean economy that make it an ideal environment for my analysis. First, local government spending varies across 222 granular spatial units. These spatial units, referred to as *districts*, partition the mainland of South Korea.[1] Each district has a local government which provisions local public goods. This local spending is financed via income tax from its residents, a part of which is locally retained while the rest is redistributed across districts.[2] Second, national tax policy reforms in 2008 and 2012 reduced the income tax rates providing a quasi-natural experiment to estimate the preferences for local government spending. Although these reforms modified national tax policies, they resulted in differential changes in local government revenues

---

[1]Districts in this paper correspond to 222 administrative units in South Korea called *Si*, *Gun*, or *Gu*. To give a sense of the scale, the total land area of South Korea is about 1 percent of the U.S. or about the same size as the state of Kentucky.

[2]The national government of South Korea redistributes local tax revenue across districts via intergovernmental transfers. The Local Subsidy Act describes a set of formulas computing the amount of intergovernmental transfers each district receives. The rules of redistribution favor districts with lower amenity values and higher population density. There was no major changes in the formula since its enactment in 1994. See Section 2.3 for more details.

because each district had a different socio-economic composition that determined tax base. Due to budget balancing, these changes in local government revenues led to equivalent changes in its spending. Third, I can observe bilateral migration and commuting decisions every 5 years from 2005 to 2015. I use restricted-access administrative data from the Population Census of South Korea to construct a geo-coded panel data set of the number of workers in terms of three locations, which I define by their district of residence 5 years ago, current district of residence, and workplace district.

To guide the analysis, I use a quantitative spatial equilibrium model with a number of features in common with Ahlfeldt et al. (2015) and Monte et al. (2018). As in their frameworks, workers decide where to work and where to live taking wages and floor-space prices into account. The model accommodates an arbitrary number of spatial units (corresponding to the districts in my data). Districts are different in terms of local amenities as a residence and in terms of productivity as a workplace, while commuting costs vary by district pair. The supply of floor space is endogenously determined for commercial or residential use. Following these frameworks, I also incorporate idiosyncratic preferences for residential and employment locations in the spirit of McFadden (1974) and Eaton and Kortum (2002).[3]

In addition to these standard features, I incorporate two key margins. First, as in Morten and Oliveira (2018), workers are heterogeneous in terms of their place of previous residence, which empirically corresponds to where people lived 5 years ago. This margin implies additional spatial frictions on top of commuting. Specifically, I allow the spatial frictions between previous and current residences (i.e., migration) and between previous residence and workplace location (i.e., job finding). Second, residential decisions depend on local government spending, which are financed through a fiscal transfer scheme that corresponds to what I observe in Korea. Local government spending may lead to agglomeration or congestion spillovers depending on the extent of rivalry associated with how much people benefit from local government goods.[4] The framework generates a gravity equation, which expresses the fraction of workers from each origin living in a residence and commuting to a workplace location as a function of spatial frictions between these locations, spatial frictions with respect to the previous residence, wages at the workplace, and government spending, density and home prices at the residence.

I implement a two-step approach to estimate people's valuation for government goods and spatial frictions. First, I recover reduced-form elasticities governing worker mobility. My identification strategy to estimate the preferences for local government spending is to compare how the decisions of migration and commuting changed due to an increase in local government spending, *ceteris paribus.* I construct instrumental variables based on the tax reforms discussed

---

[3]See Redding and Rossi-Hansberg (2017) for a review of quantitative spatial models.

[4]E.g., under full rivalry, local government goods are simply publicly provided private goods. Under no rivalry, these goods behave like an agglomeration externality.

above and historical residential density to estimate the reduced-form elasticities of worker mobility with respect to local government spending, residential density, and home prices. I find that the probability of migration increases by 1.07 percent for a 1 percent increase in local government expenditure and decreases by 0.8 percent and 0.5 percent for a 1 percent increase in residential density and home prices, respectively. The results imply that there is rivalry associated with local government goods (i.e., a dollar tax in contribution of a worker is shared with his fellow residents to a certain extent). These estimates, nonetheless, are not enough to recover the valuation of local government spending. For that, I also need to know how much people move in response to spatial frictions and wages. Using the mobility response to wages, I can then express the response to government spending in disposable-wage equivalent units and recover and estimate the utility parameters.

In the second step, I use the structural properties of the model and estimate the effects of spatial frictions (i.e., costs of migration, commuting, and job finding) and wages on worker mobility. I estimate the distance elasticities of migration, commuting, and job finding to be negative and stable over time. I show that estimating the distance elasticity of migration while not taking commuting into account and the distance elasticity of commuting without accounting for migration lead to large biases. With respect to the distance elasticity of migration, the upward bias arises because workers migrate over long distance when they are better compensated from the local labor market at the destination.[5] Estimating the distance elasticity of commuting while not accounting for migration leads to an overestimation because there are additional costs associated with commuting due to the costs of migration and job finding in addition to the direct cost of commuting explained by the distance of commuting.[6] These biases primarily arise because the previous literature studied migration and commuting under the assumption that previous residential location does not affect where they work, which the data unequivocally rejects.

Following the approach in Ahlfeldt et al. (2015), I estimate how much people move in response to wages (i.e., the Fréchet shape parameter). Using this estimate, I re-scale the estimated reduced-form elasticity of worker mobility with respect of local government spending and compute the marginal valuation of local government goods in dollars. As a result, I find that workers on average value an additional dollar of local government goods equal to 75 cents of their after-tax income. This estimate is similar to the point of estimate of Suárez-Serrato

---

[5]For example, Bryan and Morten (2018) estimates the distance elasticity of migration based on the migration patterns in the U.S. and Indonesia without taking commuting into account. Based on the migration pattern in South Korea alone, I estimate a value of the elasticity similar to theirs, which is about 4.7 times smaller in magnitude than my estimate based on both migration and commuting patterns.

[6]Ahlfeldt et al. (2015) estimates this elasticity in the context of the City of Berlin, Germany. My estimate following their estimation strategy is similar to wht they found, which is 2.1 times larger in magnitude (more negative) than the estimate based on both migration and commuting.

and Wingender (2014), who use a different source of variation in the U.S. context.[7] In order to correctly estimate how much people value local government goods, it is important to take both margins of mobility into account especially when spatial units are finely defined. The valuation of local government spending is biased upward if migration is not taken into account and downward if commuting is not taken into account.

Using the estimated general equilibrium model, I quantify the welfare consequences of the spatial distribution of local government spending. I conduct a set of counterfactual policy experiments to shed light on the optimal modes of fiscal decentralization (local taxation vs. redistribution). Across counterfactual exercises, I vary the extent of redistribution, i.e., how much local government spending depends on redistributive intergovernmental transfers relative to local taxation. Many countries around the world (e.g., Canada, Germany, Australia and Japan) make fiscal transfers across regions, similar to the South Korean system featured in this paper. I allow for counterfactual regimes to mimic what is observed in other countries, ranging from a high redistribution (as in Canada and Denmark) and no redistribution (as in the U.S.).

I find that there would be a welfare improvement if the extent of redistribution observed in 2015 were reduced. However, the complete elimination of the redistributive intergovernmental transfers would lead to a sizable loss of welfare. The results indicate that transfers of income are too high from fiscally strong districts (i.e., districts with higher average income) to the weak (i.e., districts with lower average income) under the redistribution policy observed in 2015. The benefit of the transfers in the net-receiving districts is dominated by the loss in the net-contributing districts. Lastly, I show that different assumptions on spatial mobility of workers (e.g., costless migration and prohibitively costly commuting) call for a significantly different extent of redistribution. For example, if no spatial frictions of migration and job finding are assumed as in the commuting literature, a fiscal arrangement with significantly lower redistribution appears optimal. In this scenario, a lower extent of redistribution improves the overall welfare by reducing the incentive for workers to reside in net-receiving districts at the expense of longer commute.

My paper builds upon several existing literatures. The public finance literature examines the effects of government policies on the spatial distribution of workers. Tax differentials across space incentivize workers to move across the state and country borders (Kleven et al., 2014; Akcigit et al., 2016; Moretti and Wilson, 2017). Some papers estimate positive amenity values for government spending and regulations from housing prices (Cellini et al., 2010; Black, 1999; Chay and Greenstone, 2005) in the spirit of Rosen (1979) and Roback (1982). There are only a

---

[7]Suárez-Serrato and Wingender (2014) estimate the valuation of government spending in the context of the U.S. They exploit the population revisions following the decennial Census ("Census Shock") and the measurement error in population levels during non-Census years to isolate exogenous variation in federal spending across counties in the U.S.

few papers that directly estimate how much workers value government goods using government spending. Using a spatial general equilibrium framework, Suárez-Serrato and Wingender (2014) estimates the effect of federal spending on local economies in the U.S. by exploiting changes in population levels used to determine the size of federal funding for localities due to Census shocks (Suárez-Serrato and Wingender, 2016). Fajgelbaum et al. (2019) rely on tax differences across U.S. states over time and the spatial proximity to other states to estimate worker preferences for government expenditure.[8]

My approach includes various novel features relative to these papers. First, the spatial unit used in this paper is finer than the spatial units commonly considered in the literature (e.g., states and county groups). Given the granular spatial units, I leverage both migration and commuting patterns to estimate how much workers value local government goods and services. Second, I provide a new identification strategy using national tax reforms as a source of plausibly exogenous variation in local government spending to estimate the elasticity of worker mobility. Third, I estimate the effect of residential density on worker mobility by following the standard approach used in the urban economics literature to estimate agglomeration and congestion forces (Ciccone and Hall, 1996; Combes and Gobillon, 2015; de la Roca and Puga, 2017).

This paper also contributes to the fiscal decentralization literature. The majority of the papers in this literature focus on theoretically and empirically examining the consequences of the changes in fiscal autonomy of local governing entities.[9] There are relatively few empirical papers studying the effects of policy instruments employed for fiscal decentralization (e.g., local taxation and redistribution). Government goods and services are often public, thus creating fiscal spillovers. Wildasin (1980) finds that households may locate in an optimal fashion in the presence of the spatial distribution of local government spending and notes that fiscal spillovers may result in non-optimality. Fajgelbaum and Gaubert (2018) characterize the optimal transfers for efficient allocations and the policies implementing the transfers. Albouy (2012) presents a theoretical framework to determine efficient and equitable transfers across localities and evaluates the welfare consequences of the equalization policy in Canada. I contribute to the literature on fiscal decentralization by computing the optimal mix of location taxation and redistribution.

Lastly, this paper contributes to a growing literature on quantitative economic geography models. There are a number of recent papers that have studied the migration and commuting

---

[8]Gelbach (2004) focuses on the female population in the U.S. eligible for state welfare programs and finds that the interstate migration patterns of this population are not sensitive to the distribution of welfare benefits across states.

[9]For instance, Fisman and Gatti (2002) documents that fiscal decentralization leads to a lower level of corruption. Bianchi et al. (2019) show that fiscal decentralization led to a higher female labor force participation because local governing authorities expanded nursery schools. See Oates (1999) for a broader literature review on fiscal federalism.

decision, separately. In the case of migration, Bryan and Morten (2018) study the cost of migration as a source of friction that results in labor market misallocation using the case of Indonesia. Morten and Oliveira (2018) quantify the impact of transport networks using the construction of a radial highway system in Brazil when workers can migrate across space. Moretti and Wilson (2017) estimate the negative effects of tax rate differences across states on the migration of star-scientists in the U.S.[10] In the case of commuting, Ahlfeldt et al. (2015), and Tsivanidis (2019) study the commuting patterns and their contributions to the spatial distribution of economic activity in the city of Berlin, Germany and in the city of Bogotá, Columbia, respectively. The literature on migration assumes that workers live and work in the same locations. The literature on commuting assumes often implicitly zero spatial frictions associated with migration and job finding. Monte et al. (2018) have a notion of migration in addition to commuting; however, they assume that where workers migrate from does not affect their commuting decisions. To the best of my knowledge, this paper is the first to present a spatial equilibrium model featuring both bilateral migration and commuting in the economic geography literature. Furthermore, these papers concerning the geographical mobility of workers do not study the roles of the public sector.

The remainder of this paper is structured as follows. In Section 2, I describe the data sources and the key aspects of the South Korean economy. In Section 3, I present a partial equilibrium model in which workers choose where to live and where to work in the presence of local government goods and services as well as costs associated with mobility. Then, I estimate the elasticities of worker mobility with respect to local government spending, residential density, and home prices in Section 4. Section 5 focuses on the effects of spatial frictions on worker mobility and estimate three reduced-form elasticities measuring the responsiveness of worker mobility (migration, commuting, and job finding) with respect to distance between localities. In Section 6 and 7, I embed the partial equilibrium model presented in Section 3 into a general equilibrium setup and describe how I parameterize the model. Finally, I consider counterfactual policy experiments concerning the extent of redistribution and its aggregate welfare implications in Section 8. Section 9 concludes.

2   Data and Background

In this section, I discuss some key aspects of the South Korean economy and the data I have collected to study how the spatial distribution of local government spending affects the spatial mobility of workers and, more broadly, how it affects the aggregate welfare of workers. Specifi-

---

[10]There are more papers studying the migration patterns in the U.S., Vietnam, and Brazil based on spatial equilibrium models: e.g., Piyapromdee (2017); Albert and Monras (2019); Balboni (2019); Pellegrina and Sotelo (2019). My model abstracts away from the dynamic model presented in Caliendo et al. (2019).

cally, in Section 2.1, I discuss main data sources of the key variables for my empirical study. In Section 2.2, I define the geographic units used in this study and document the migration and commuting patterns in South Korea. Lastly, Section 2.3 discusses the national policies on local public finance and describe national tax reforms in 2008 and 2012.

2.1 Data

The observed spatial distribution of workers is a consequence of decisions on two margins of geographical mobility of workers—migration and commuting. Therefore, my empirical analysis has a specific data requirement. First, I need data that records worker's previous residence, current residence, and current employment location. Second, the data has to be spatially representative. Third, local government spending should vary at the same spatial disaggregation across which workers activley make both migration and commuting decisions. South Korea is one of the few countries, which meet all of the requirements.

My main data source for the spatial distribution of workers is the restricted Population Census of South Korea. The Population Census of South Korea is conducted every five years and sample 20 percent of the entire population. I use the three most recent waves from 2005, 2010, and 2015. I restrict the sample to working male household heads between the ages of 25 to 60, who commute a round trip of less than 180 kilometers.[11] The sample size is about 3.5 million households. The Census questionnaire asks district of residency five year ago, current district of residence, and district of workplace location. Based on this information, I construct a panel data set of the distribution of workers by residence five years ago, current residence, and workplace location.

Data on local government spending was collected from the administrative data (Yearbook of Local Public finance) from the Ministry of Interior and Safety of South Korea. I digitized local government information by total revenue and revenue from the following sources: local income taxes and intergovernmental transfers. This information allows me to recover the share of the intergovernmental transfer each locality received from the national government in a given year. In addition, the Ministry of Land, Infrastructure, and Transport publishes the land price fluctuation rates for each district. I collected this information for 2005, 2010, and 2015. The fluctuation rate is defined as a ratio of the average land price in a given year to the average land price in 2004.

I supplement the main data set with local characteristics in 2015 using the administrative

---

[11]The reason for restricting the same to male household heads is motivated by the fact that migration decisions are made at the household level. Over 90 percent of the households in the Population Census have male household heads. The female labor force participation in South Korea is one of the lowest among the OECD countries (Lee, 2017). The age restriction is to only include workers who have completed education. Also, less than 1 percent of workers report a commuting distance over 90 kilometers in each direction.

data from various government agencies to complete the parameterization of the spatial general equilibrium model I present later in the paper. The two key variables are wages and housing prices.[12] A major limitation of the Population Census is that the information on wealth and income is not surveyed. Instead, I use the Economic Census of 2015, which surveys the universe of establishments, and compute the average annual wages in each district. The Ministry of Land, Infrastructure, and Transport maintains the universe of housing transactions from 2006 to 2015. I construct district-level prices per unit of floor space in 2015 by employing a Case-Shiller type repeated sales approach at the district level, similarly done in Ahlfeldt et al. (2015). Lastly, I compute distance between every pair of districts by connecting their centroids.[13]

2.2   Spatial Mobility in South Korea

South Korea has the 11th largest economy in the world, comparable to Canada and Spain in terms of GDP and GDP per capita, respectively. While South Korea is only about 1 percent of the U.S. geographically, the population level was as high as 51 million in 2015, about 16 percent of the population in the U.S.

I describe two dimensions of spatial mobility—commuting and migration—in South Korea. First, I begin with the commuting patterns in South Korea. Workers in South Korea spend about 7.3 percent of their workday commuting between their residence and workplace locations, reflecting the commuting patterns documented in Monte et al. (2018) for the U.S. and Schafer (2000) for 26 countries around the world.[15]

In Panel A of Table 1, I report summary statistics on commuting patterns in 2005, 2010, and 2015. On average, about 27 percent of residents work outside their district of residence and about 30 percent of workers commute to work from other districts. In addition, I plot the fraction of residents commuting to other districts against distance between residence and

---

[12]See Appendix A for the complete list of additional variables and their sources.

[13]In addition to the key variables explained above, I collect other local characteristics (e.g., land use, suicide rates, divorce rates, and number of firms) for cross-validation exercises and over-identification checks carried out later in the paper. See Appendix A for details.

[14]As of 2005, there are 226 local governments, four of which were split as a consequence of redistricting. I keep the administrative units consistent to the administrative boundaries set in 2005 by defining groups of districts for those which underwent redistricting.

[15]See Redding and Turner (2015) for further discussion on cost of commuting and transportation costs.

workplace in Panel (a) of Figure 1. I observe that the probabilities of commuting decreases in distance. This implies that the cost associated with commuting increases in distance consistent with prior literature (Ahlfeldt et al., 2015; Monte et al., 2018; Tsivanidis, 2019).

Second, with respect to migration, about one in seven households migrate across district borders annually; the implied annual inter-district migration rate is around 13 percent.[16] Aggregated at the province level (i.e., 16 groups of districts), the annual migration rate in South Korea is 5 percent, similar to the inter-county migration rate in the U.S. reported in Molloy et al. (2011).[17] In Panel B of Table 1, I report the probabilities of migration over 5 years at the district level. On average, about 19 percent of residents in a district are migrants who have migrated from other districts within 5 years, while 18 percent of residents have migrated out of their residence in the past 5 years. Panel (b) of Figure 1 plots the probabilities of migration conditional on location of origin against the distance between the origin and current residence. In line with the literature on migration, I also observe that the probability of migration decreases with distance (Bryan and Morten, 2018; Morten and Oliveira, 2018).

In Figure 2, I plot the number of households in the figure on the left and local expenditure on the right by district. There are districts with generous local expenditure and many households. This pattern suggests that workers are more likely to reside in districts with more generous provision of local government goods and services. Additionally, local determinants like wages, home prices, and amenities also influence worker's migration and commuting decisions. The spatial distributions of these additional forces also explains the distribution of workers across localities. Figure 3 provide suggestive evidence that workers are willing to migrate further and commute longer to live in a district with a relatively higher level of local government expenditure and lower home prices.

2.3   Local Government Revenue and Tax Reforms in 2008 and 2012

The total local government expenditure accounts about 8 percent of South Korean GDP in 2015. I focus on two main sources of local government revenue: local income taxes and intergovernmental transfers, which constitute 14 percent and 72 percent of local government spending, respectively.[18] Panel C of Table 1 reports the summary statistics of local government expenditure for 222 districts in the years 2005, 2010, and 2015. The average total local expenditure is

---

[16]I do not observe annual migration patterns in the Population Census. Instead, the annual migration rates are calculated using the restricted-use administrative records of the universe of migrants in South Korea during the same time period (2005-2015). The migrant records are not used for analysis in this paper because it does not provide information on where migrants work.

[17]Therefore, the contrast in annual migration rates between South Korea (inter-district) and the U.S. (inter-county) can be explained by geographical size differences between U.S. counties and South Korean districts.

[18]The remaining 14 percent of local government revenue is comprised of non-tax receipts (e.g., fees, charges, and fines) and borrowing, last of which is only about 0.06 percent on average. Hereon, I refer to the sum of local income taxes and intergovernmental transfers as local government revenue or expenditure.

363 billion KRW (approximately 363 million USD); the average per-capita local expenditure is 7,638 USD, widely ranging from 906 USD to 29,622 USD.[19]

There is substantial spatial variability in the degree to which districts depend on local income taxes and intergovernmental transfers for their total local spending. Local income tax constitutes on average about 16.8 percent of the total local revenue. Intergovernmental transfers constitutes about 83.2 percent of total local spending.[20] The share of local government revenue from local income tax ranges from 2.1 percent to 56 percent.

The national fiscal policies (progressive income tax system and extents of fiscal decentralization and redistribution) and local tax bases (number of workers and their income) determine local government revenue. The Local Autonomy Act—first enacted in 1949—was revived in 1991 after 30 years of suspension due to military dictatorships that ended in 1987. The purpose of the Act was to "strive for democracy and efficiency of local autonomous administration and to ensure balanced development of local areas..." (Local Autonomy Act, 1991). The national government amended the Local Tax Act and Local Subsidy Act to enable local autonomy in 1994. The Local Tax Act and Local Subsidy Act together with the Income Tax Act promulgates progressive income tax rates. I will refer to the collection of these three Acts as the national fiscal policies.

The national fiscal policies determine the size of local governments in two ways. First, local governments collect income tax from their residents according to the income tax rates outlined in the Income Tax Act, which is uniform across all districts. Local governments retain a fixed share of their income taxes and deliver the rest to the national government. I refer to the fixed share as *local-national revenue sharing* and the amount of income tax revenue left at the local level as *local income tax.* Lastly, the national government allocates a fixed share of its tax revenue for redistribution. Then, the national government makes *intergovernmental transfers* to each local government, calculated by a set of formula determining the shares of the total fund allotted to each local government.[21] In sum, the extent of fiscal decentralization (the fraction

---

[19]For simplicity, I will continue assuming the unit of government spending (and wages) in USD throughout the remainder of the paper.

[20]In Appendix B.1, Figure B.1 plots the spatial distribution of local tax revenue in Panel (a) and the ratios of local tax revenue to total local spending in Panel (b). Likewise, Panel (c) and (d) plots the spatial distribution of intergovernmental transfers and the contribution of intergovernmental transfers to total local spending for each district. According to Figure B.1, there is a considerable variation in how fiscally strong localities are.

[21]The Local Subsidy Act details the formula employed to determine how much intergovernmental transfers to be rebated to each locality. The overarching objective of intergovernmental transfers is to help develop "the public administration of local governments in a sound manner with the adjustment of their finances by subsidizing financial resources necessary for the public administration of local governments" (Local Subsidy Act, 1994). There are a number of countries both developed and developing (e.g., Germany, UK, Canada, Australia, and India) with a similar local finance instrument (equalization grants) to promote balanced financial capacities horizontally. While the U.S. does not have a federal system directly aiming to reduce differences in fiscal capacities across localities, many of the federal grants and policies have features that are implicitly equalizing across states and localities (e.g., EITC, SNAP, medicare, and medicaid).

of total tax revenue local governments spend) and the rules of redistribution (how to allocate intergovernmental transfers across local governments) determine local government revenues.

Tax Reforms

there were two major reforms on national tax policies: one in 2008 and the other in 2012. Figure 4 plots the marginal income tax rates before and after the tax reforms in 2008 and 2012. In 2008, the Income Tax Act was amended to substantially decrease income tax rates across income brackets: from 11 percent to 8.8 percent for the low income bracket (annual income less than 12 million KRW); from 22 percent to 18.7 percent for the middle income group (12 million to 46 million KRW); and from 33 percent to 28.2 percent for the high income group (46 million to 88 million KRW).[22] In 2012, the national government further reduced the income tax rates to 6.6 percent for the low income group, to 16.5 percent for the middle, and to 26.4 percent for the high group. The tax reforms did not affect the rules of redistribution outlined in the Local Subsidy Act. Figure B.2 in Appendix B.1 plots the current shares of intergovernmental transfers to the shares 5 years ago. The estimated slopes comparing the redistribution policies in a given year to these five years ago are close to 1.

3   Discrete Choice Model of Worker Location Decisions

In this section, I present a discrete choice model, in which workers make decisions on migration and commuting. In the model, a worker decides where to live and where to work, taking wages, prices of residential floor space, local government goods and services, and the location choices of all other workers into account. My model is different from the spatial equilibrium models commonly used in the recent literature examining the spatial mobility of workers in two ways. First, similar to Fajgelbaum et al. (2019), I augment the model by introducing goods and services provisioned by local governments. Second, the model features both commuting (Ahlfeldt et al., 2015; Tsivanidis, 2019) and migration decisions (Bryan and Morten, 2018; Morten and Oliveira, 2018) which have been independently studied. There are iceberg costs of worker mobility rising from three spatial frictions: migration, commuting, and job finding. The key prediction of the model is a gravity equation which summarizes the distribution of workers in terms of initial residence, current residence, and workplace location.

---

[22]The total number of income brackets had been four until the second amendment in 2012, which introduced one additional income brackets for the even richer. For my analysis, I focus on the lowest three income brackets which include more than 95 percent of workers in South Korea according to the Ministry of Strategy and Finance of South Korea. I also note that the first reform in 2008 resulted in small changes in the cutoffs of each bracket to account for inflation since the last change in 1994. Since the first reform, the cutoffs for the lowest three income brackets remain the same.

3.1 Model Environment

The whole economy has R measure of workers (also, interchangeably referred to as residents) and comprises of J discrete number of spatial units (i.e., districts), indexed by r for current residence, m for workplace, and o for initial residence. As a residence, each district is characterized by exogenous local amenities B_r, per-unit floor space price Q_r, and local government goods and services g_r. As a workplace, a district is characterized by wage w_m, which is subject to income tax. Therefore, workers commuting to district m receive after-tax income equal to $(1 - \tau_m)w_m$ for their private consumption of the single final good c_{rm} and residential floor space h_{rm} in residence of district r. In addition, there are iceberg costs of worker mobility across space in three dimensions: migration D_{or}, commuting D_{rm}, and job finding D_{om}, summarized in a single disutility index $D_{orm} = \varepsilon_{orm} D_{or} D_{rm} D_{om}$ where ε_{orm} is a stochastic error term following a log-normal distribution with its mean equal to 1. The first two spatial frictions follow the standard formulations in the literature on migration and commuting. The spatial friction captures the iceberg cost associated with finding a job in workplace location m from initial residence o.

Workers are born in or assigned to initial residence o (also, interchangeably referred to as origin). The initial distribution of workers is given by π_o. I assume each worker inelastically supplies one unit of labor. After observing an idiosyncratic utility shock for every possible pair of residence r and workplace m, a worker chooses a residence-workplace pair that maximizes his utility given his initial residence, after-tax wages, floor space prices, local government expenditure, location choices of other workers, local amenities, and the iceberg costs of mobility.

3.2 Worker's Location Decisions

The preferences of a worker i are defined over amenities, consumption of the single final good, consumption of floor space for housing, public goods, and the iceberg costs associated with migration, commuting, and job finding. The direct utility of a worker i who chooses to move from his initial residence o to a new residence r and commutes to a workplace m is $z_{irm} u_{orm}(c_{rm}, h_{rm})$, where c_{rm} is consumption of the single final good (numeraire) and h_{rm} is consumption of floor space for housing.

First, $u_{orm}(c_{rm}, h_{rm})$ corresponds to the systemic component of the preference and follows the Cobb-Douglas form:

$$u_{orm}(c_{rm}, h_{rm}) = \frac{B_r}{D_{orm}} \left(\frac{c_{rm}}{\beta} \right)^\beta \left(\frac{h_{rm}}{1 - \beta} \right)^{1-\beta} g_r^\lambda. \tag{1}$$

Amenity fundamental B_r captures intrinsic residential characteristics that make district r more or less attractive to live in (e.g., the weather, beaches, and scenic views). The parameter β

determines the share of expenditure on the final consumption good.[23] Following Fajgelbaum et al. (2019) and Albouy (2012), real government expenditure enjoyed by each worker living in district r g_r is local government expenditure G_r, normalized by a function of the total number of workers living in district r R_r:

$$g_r = \frac{G_r}{R_r^\theta}. \tag{2}$$

The parameter θ controls the extent to which local government goods and services are rival and ranges from 0 if non-rival (pure public good) to 1 if rival (publicly-provided private good). The parameter $\lambda \geq 0$ captures the weight of local government goods and services in preferences relative to consumption of the single final good and floor space for housing.

Given unit price of floor space for housing Q_r and after-tax wage $(1 - \tau_m)w_m$, the budget constraint is $c_{rm} + Q_r h_{rm} = (1 - \tau_m)w_m$. The Cobb-Douglas preference implies that β share of after-tax wage is allocated to the consumption of the single final good and the rest to the consumption of residential floor space. Therefore, the indirect utility of a worker i choosing to live in district r and commute to district m is $V_{iorm} = z_{irm}v_{orm}$, where

$$v_{orm} = \frac{B_r(1 - \tau_m)w_m}{D_{orm}Q_r^{1-\beta}}\left(\frac{G_r}{R_r^\theta}\right)^\lambda. \tag{3}$$

The systemic component of the indirect utility increases in local amenities B_r, wage w_m, and local government expenditure G_r, while it decreases in per-unit price of floor space Q_r, residential population R_r, and spatial frictions $D_{orm} = \varepsilon_{orm}D_{or}D_{rm}D_{om}$. The composite iceberg cost associated with migration, commuting, and job finding D_{orm} enters the indirect utility function multiplicatively. Therefore, there is an isomorphic formulation in which after-tax wages are reduced due to commuting and job finding costs, and amenity values are decreased due to migration cost.

Second, z_{irm} is an idiosyncratic preference shock that captures the idea that each individual worker has idiosyncratic reasons to find a residence and a workplace more or less attractive. I model this heterogeneity in preference in spirit of McFadden (1974) and Eaton and Kortum (2002). For a worker i residing in district r and working in workplace location m, the idiosyncratic component of his utility is drawn from an independent Fréchet distribution:

$$\Pr(z_{irm} < z) = \exp(-T_r M_m z^\epsilon), \tag{4}$$

where the parameter $T_r > 0$ determines the average utility of living in district r; the parameter $M_m > 0$ determines the average utility of working in district m; and the shape parameter $\epsilon > 1$

---

[23]Davis and Ortalo-Magné (2011) provide empirical evidence supporting the constant housing expenditure share, using the U.S. as a case study.

governs the dispersion of the utility draw.[24] Then, the distribution of workers living in district r and working in district m by initial residence o:

$$\pi_{orm} = \frac{\left(\frac{\tilde{B}_r(1-\tau_m)\tilde{w}_m}{D_{orm}Q_r^{1-\beta}}\left(\frac{G_r}{R_r^\theta}\right)^\lambda\right)^\epsilon \pi_o}{\sum_{r'=1}^J \sum_{m'=1}^J \left(\frac{\tilde{B}_{r'}(1-\tau_{m'})\tilde{w}_{m'}}{D_{or'm'}Q_{r'}^{1-\beta}}\left(\frac{G_{r'}}{R_{r'}^\theta}\right)^\lambda\right)^\epsilon} \equiv \frac{\Phi_{orm}\pi_o}{\Phi_o}, \text{where } \Phi_o = \sum_{r'=1}^J \sum_{m'=1}^J \Phi_{orm}. \quad (5)$$

Because some of the unobserved local characteristics (i.e., T_r and M_m) always appear in the gravity equation together with unobserved local amenities B_r and wages w_m, I define the following composite terms denoted by a tilde: adjusted amenities $\tilde{B}_r = B_r T_r^{1/\epsilon}$ and adjusted wages $\tilde{w}_m = w_m M_m^{1/\epsilon}$.

Workers are more likely to live in a residences with a high amenity value and local government expenditure and lower per-unit floor space price, net of congestion/agglomeration forces and migration costs.[25] Workers are more likely to commute to workplace locations with higher after-tax wages net of commuting and job finding costs.

4   Key Reduced-Form Elasticities of Worker Mobility

In this section, I estimate the reduced-form elasticities of worker mobility with respect to local government expenditure, residential density, and floor space prices derived from the observed distribution of worker mobility in South Korea. Section 4.1 discusses an econometric specification, which I derive using the gravity equation, a key prediction of the spatial equilibrium model presented in the preceding section. In order to consistently estimate the reduced form elasticities of interest, I exploit the episodes of national tax reforms discussed in Section 2.3 as well as information on the historical residential density. In Section 4.2 and 4.3, I present the estimation results and discuss the interpretation and robustness of the estimated reduced-form elasticities.

---

[24]The indirect utility V_{irmr_0} is Fréchet distributed since V_{iorm} is a monotonic function of the Fréchet distributed idiosyncratic preference shock z_{irm}. The maximum utility is itself Fréchet distributed appealing to the stability postulate.

[25]In Appendix D.1, I discuss how I derive the gravity equation (5). It is also general enough to produce the gravity equations summarizing the spatial distribution of workers that the literature on commuting and migration have considered based on economic geography models (Ahlfeldt et al., 2015; Bryan and Morten, 2018; Morten and Oliveira, 2018; Monte et al., 2018; Moretti and Wilson, 2017). In Appendix D.2, I show that the gravity equations used elsewhere can be derived based on the gravity equation (5).

4.1 Estimation Strategy

The gravity equation (5) describes how workers sort across districts in terms of residential and workplace locations from previous residence. I take the log transformation of both sides of the gravity equation and obtain the following econometric specification by augmenting the terms with time subscript whenever applicable to permit the panel structure of the data:

$$\ln \pi_{orm,t} = \phi_{om,t} + \phi_{or} + \phi_{rm} + \underbrace{\lambda\epsilon}_{\beta_G} \ln G_{r,t} - \underbrace{\theta\lambda\epsilon}_{\beta_R} \ln R_{r,t} - \underbrace{(1-\beta)\epsilon}_{\beta_Q} \ln Q_{r,t} + \zeta_{orm,t}. \qquad (6)$$

The coefficients in front of log local government expenditure ($\beta_G = \lambda\epsilon$), log number of workers living in r ($\beta_R = \theta\lambda\epsilon$), and log prices of floor space ($\beta_Q = (1-\beta)\epsilon$) are the reduced-form elasticities in question and are functions of structural parameters. The job finding fixed effects interacted with year dummy variables $\phi_{om,t}$ flexibly capture the workplace-specific factors (e.g., after-tax wages and average utility from working in district m) and the factors specific to the origins (e.g., number of workers who used to live in o and the denominator of the gravity equation (5)) as well as the iceberg cost of job finding. The migration fixed effects ϕ_{or} and the commuting fixed effects ϕ_{rm} capture the time-invariant component of the iceberg costs of migration and commuting as well as the intrinsic residential characteristics of district r that makes it a more or less attractive place to live in.[26] Lastly, the error term $\zeta_{orm,t}$ includes the rest of the factors in equation 5 (i.e., adjusted amenities and time-varying stochastic components of the spatial frictions net of costs associated with job finding).

The errors in Equation 6 can be correlated in two ways. First, there is a classic clustering concern explained in Moulton (1990). Second, one may worry about the serial correlation over time within a panel dimension Bertrand et al. (2004). In order to address these concerns, I report standard errors that are robust to heteroskedasticity and allow multi-way clusterings. I allow errors to correlate across previous residences and across workplace locations sharing the same current residence in a given year. In addition, the serial correlation within each of the panel dimension (a triplet of previous residence, current residence, and workplace location) over time.

4.1.1 Fixed Effects

The mapping between the econometric specification (6) and the gravity equation from the spatial model (5) helps to understand potential confounders and consequent biases. First, the job

---

[26]Note that land area for each district is absorbed into the migration and commuting fixed effects because area is a time-variant feature of each locality. This implies that β_R can be interpreted as the elasticity of worker's mobility with respect to residential density.

finding fixed effects interacted with year dummy variables $\phi_{om,t} = \ln(1-\tau_{m,t})^\epsilon \tilde{w}_{m,t}^\epsilon \exp(-\delta\epsilon d_{om})\pi_{o,t}/\Phi_{o,t}$ control for the benefit from choosing to work in m net of the job finding cost from previous residence o. Workers are more likely to choose workplaces with higher net benefits. Given higher returns from workplace location, workers are willing to accept a lower amount of local government spending at their residential location. Furthermore, worker's valuation of a given workplace location depends on their origin because for example they rely on their network to find higher paying jobs, and this network is usually formed at the origin (Card, 2001; Cadena and Kovak, 2016). Thus, if one does not control for the different levels of the attractiveness of the nearby workplace by origin, the OLS estimate of β_G will be downward biased.

Next, a higher net labor market return attracts residents. This positive correlation between the residential density and the labor market return biases the OLS estimate of β_R upward. Workers with higher after-tax wages would be able to afford higher housing prices. Similarly, excluding the job finding by year fixed effects biases the OLS estimates of β_R and β_Q because residential density and home prices partially reflect the fact that there are attractive workplaces nearby for workers of a given origin.

Second, omitting the migration fixed effects $\phi_{or} = -\rho\epsilon d_{or}$ and commuting fixed effects $\phi_{rm} = -\kappa\epsilon d_{rm}$ are likely to bias the OLS estimate of β_G downward. While the costs of migration and commuting inhibit worker mobility, workers may choose residences with higher local government expenditures to offset their migration and commuting costs. Districts that are attractive to live in are likely to have higher housing prices and residential densities. If so, the costs of migration and commuting are likely correlated positively with the residential density and housing prices, again in the sense of compensating differentials. Then, OLS estimates of β_R and β_Q would be biased downward.[27]

4.1.2 Endogeneity

Even after conditioning on the set of fixed effects discussed above, OLS estimates of β_G, β_R, and β_Q suffer from endogeneity due to omitted variable bias and measurement errors. The error term $\zeta_{orm,t} = \ln \tilde{B}_{r,t}^\epsilon \varepsilon_{orm,t}^{-\epsilon}$ includes the adjusted local amenity values. Local government expenditures and local amenities are likely negatively correlated because redistributive intergovernmental transfers favor places with low amenity values *ceteris paribus*. This negative correlation between amenity values and government expenditures would generate a downward bias in the OLS

---

[27]Note that local government expenditure, residential density, and housing prices are correlated with each other. It is useful to have a sense of the potential directions of bias in the conventional way by thinking about the relationship between omitted variables and the dependent variable and the relationship between omitted variables and the endogenous regressors. Nevertheless, the covariances among the endogenous variables as well as their relationship with an omitted variable need to be taken into account in order to properly characterize the directions of potential omitted variable bias. The system of equations summarized in (38) in Appendix shows the complexity of the omitted variable bias with three endogenous variables.

estimate of β_G. Next, districts with higher amenities attract inflows of migrants, which lead to a higher residential population. This means the OLS estimate of β_R would be overestimated. Lastly, high amenity values would be priced into home prices in the sense of hedonic pricing. Then, the OLS estimate of β_Q would suffer from an upward bias in this case toward zero.

Furthermore, there is an additional concern of measurement error with respect to Q_r. I do not directly observe the home prices for 2005, 2010, and 2015. Instead, I use data on land prices as a proxy. Assuming classical measurement error, an OLS estimate of coefficient β_Q would be attenuated. In fact, because all of the endogenous regressors are correlated with each other, all the other OLS estimates would also be biased.

4.1.3 Instrumental Variables

Because the estimating equation (6) has three endogenous variables, in order to consistently estimate their coefficients, I propose three instrumental variables based on the national tax reforms and the historical values of residential density. For each district r, I first construct two instrumental variables, exploiting the episodes of tax reforms in 2008 and 2012 discussed in Section 2.3:

$$IV_{r,t}^b = \tau_{b,t}\pi_{b|r,2000}, \tag{7}$$

where the income tax rates $\tau_{b,t}$ change over time; subscript b denote each of the two income brackets I use (low and high). The values of $\tau_{b,t}$ are unique in each year t and income bracket b because the reforms took place between the years when the Population Census was conducted (2005, 2010, and 2015). Furthermore, I leverage the variation in the pre-determined share of workers by educational attainment level b, symmetrically defined in terms of two levels (low for workers who have completed high school at most and high for workers with some college degrees), for each district in 2000 $\pi_{b|r,2000}$. The predetermined local educational distribution proxies the distribution of workers by income brackets in each district.[28]

The instrumental variables ($IV_{r,t}^{low}$ and $IV_{r,t}^{high}$) capture the tax contributions of low and high income groups predicted by income distribution in 2000. Therefore, by construction, the relevance of the instrumental variables follows immediately from the local government budgetary structure: government expenditures increase in tax contributions. To satisfy the exclusion restriction, the instrumental variables must not directly influence workers to prefer one residence over another, except through their impacts on local government expenditures, floor space prices, and residential densities. There are two sources of variation in the proposed instruments. One

---

[28]The sum of the proposed instrumental variables ($\sum_b IV_{r,t}^b$) shares the same structure as Bartik-type instruments widely used on the literature (Bartik, 1991). Borusyak et al. (2018) and Goldsmith-Pinkham et al. (2018) discuss sources of variation in shift-share instruments and identification approaches. See Adao et al. (Forthcoming) for inference procedures when employing shift-share instruments.

source is tax rate changes over time. Conditional on wages ($\phi_{om,t}$), workers are subject to the same tax rates regardless of their residential and employment locations. Thus, the tax rates do not directly affect their location decisions. Another source is the cross-sectional variation in the educational distribution within each district in 2000. Although my model does not take a stance on the sorting by skill levels, the previous literature has found that workers sort based on education or skill levels as skill-mix determines residential amenities (Eeckhout et al., 2014; Diamond, 2016; Fajgelbaum and Gaubert, 2018). The validity of the proposed instruments still holds as long as the tax reforms changed the educational composition in each district. In Appendix B.6, I confirm that the tax reforms are orthogonal to changes in educational composition within each district over time. Therefore, the instrumental variables constructed based on the national fiscal policy reforms remain valid with potential sorting by skill levels.

Second, the last instrumental variable $IV_{r,t}^R$ is based on the historical residential density as previously used in Ciccone and Hall (1996) and de la Roca and Puga (2017). Specifically, I use the natural logarithm of the number of households in r thirty years ago ($\ln R_{r,t-30}$) as the data allows a lag of up to 30 years: the number of households in 1975, 1980, and 1985. As Combes and Gobillon (2015) explain, historical values of residential density are usually considered relevant due to inertia in local population as local housing stock and infrastructure last over time. They are also believed to be exogenous to contemporaneous local characeristics that affect worker mobility because the changes in the type of economic activity and historical events like war reshape the economic landscape.[29] The validity of the instrument hinges on the assumption that historical residential densities do not directly affect the worker location decisions today.[30] This assumption is violated in the unlikely situation in which workers rely on the population levels 30 years ago, instead of its contemporaneous or more recent levels, when deciding where to live today.

To consistently estimate the reduced-form elasticities of worker mobility with respect to local government expenditure, residential density, and home prices (β_G, β_R and β_Q), I use the two-stage least squares (2SLS) estimator with the following identification assumption:

$$
E \left[\begin{array}{ll} IV_{r,t}^{low}\zeta_{orm,t} & |\phi_{om,t},\ \phi_{or},\ \phi_{rm} \\ IV_{r,t}^{high}\zeta_{orm,t} & |\phi_{om,t},\ \phi_{or},\ \phi_{rm} \\ IV_{r,t}^{R}\zeta_{orm,t} & |\phi_{om,t},\ \phi_{or},\ \phi_{rm} \end{array} \right] = 0. \tag{8}
$$

---

[29]In the case of South Korea, a series of military dictatorship lasted about three decades until 1987.

[30]The validity of the historical residential density as an instrumental variable can be also justified using the demographic balancing equation used in demography (Preston et al., 2000).

4.2 Estimation Results

In Table 2, I report the OLS estimates of the elasticities of mobility to local government expenditure, residential density, and home prices. In Column (1), I report the OLS estimates without including any fixed effects. The OLS estimate of β_G is negative, against the expectation that workers value local government goods. The estimated coefficient in front of the log number of households is 0.12, which implies strong agglomeration. According to the estimated coefficient of β_Q, a 1 percent increase in home prices decreases worker mobility by 0.042 percent.

In Column (2), I report the OLS estimates with the fixed effects of job finding interacted with year dummy variables. Compared to the estimate in Column (1), the OLS estimate of β_G increases to 0.097. This increase can be explained by netting out the negative correlation between local government expenditures and labor market returns from redistributive intergovernmental transfers. Furthermore, the estimated elasticity of worker mobility with respect to residential density decreases to 0.061, implying that there is a positive association between after-tax wages discounted by the cost associated with job finding and residential density in line with intuition. Lastly, the OLS estimate of β_Q increases to -0.01; however, this estimate is statistically not different from zero. On the one hand, the increase in the estimate of β_Q is against the direction of bias associated with omitting the fixed effects. On the other hand, the estimated value is likely a result of an attenuation bias due to measurement error.

In Column (3) and (4), I gradually add the fixed effects of migration pairs and commuting pairs to purge out the confounding effects of costs associated with migration and commuting on worker mobility. Because compensating differentials imply a positive correlation between the costs of mobility and local government expenditures, the coefficient estimate of β_G should increase as a result of the additional fixed effects. However, the OLS estimate of β_G changes little in Column (3) and (4). The result reflects the omitted variable bias towards zero from unobserved local amenity values, which are negatively correlated with local government expenditures and positively affects worker mobility. With respect to the estimates of β_R and β_Q, the OLS estimates increase compared to the estimated values reported in Column (2) in line with the potential directions of bias discussed.[31]

Table 3 summarizes the two-stage least squares results in Column (2), (3), and (4) and compares the results with the OLS results in Column (1). Note that the OLS estimates reported in Column (1) are the same as the ones in Column (4) of Table 2. First, according to the estimates

---

[31]The estimates in Column (4) is based on the fully saturated specification (6). According the estimated coefficients in Column (4), worker mobility increases by 0.1 percent with respect to 1 percent increase in local government expenditure and by 0.59 percent with respect to 1 percent increase in residential density. The estimated elasticity of worker mobility with respect to home prices is not only statistically insignificant, but also economically small. The OLS estimates are contaminated by measurement errors in home prices and the omitted variable bias from excluding local amenity values that make residence more attractive and are correlated with the included regressors.

in Column (2), log local government expenditure is positively correlated with the predicted tax contributions from the low and high income groups, IV^{low} and IV^{high}. The magnitudes of the estimates are similar because both tax contributions are measured in KRW. The lag residential density IV^{R} is also positively correlated with log local government expenditure. Second, the current residential density is positively correlated with the predicted tax contributions, but negatively correlated with the historical residential density. Conditional on the set of fixed effects, a negative coefficient in front of the historical residential density implies that the districts that grew at higher rates 30 years ago currently grow relatively slower. The last first stage result concerns the home prices. Log home prices are positively correlated with the predicted tax contributions as well as the historical residential density. All the the coefficients reported in Column (2), (3), and (4) are statistically different from zero at the 1 percent level. To formally test the strength of the first stage results, I compute various F-stats including SW conditional F-stats, which test the explanatory power of the excluded instruments in the presence of multiple endogenous variables (Sanderson and Windmeijer, 2016; Stock et al., 2002). I report the SW conditional F-stats and verify the strength of the first stages.

In Column (5) of Table 3, I report the 2SLS estimates of the elasticities of worker's mobility to local government expenditure, residential density, and home prices. First, the estimated elasticity to local government expenditure is statistically different from zero and substantially larger compared to the OLS estimate in Column (1). As discussed earlier, this large increase implies that there is a substantial downward bias rising from omitting time-varying local amenities, which are negatively correlated with local government expenditures, but make residences more attractive. The result indicates that one percent increase in local government expenditure increases the probability of worker's mobility (equivalently the conditional probability of migration) by 1.07 percent.

Second, the estimated elasticity of worker mobility with respect to residential density becomes negative. However, I cannot reject the null hypothesis that the estimate is different from zero. Statistical insignificance notwithstanding, the change in the sign of the elasticity indicates that there is a considerable bias toward zero resulting from omitting local amenities which the 2SLS strategy addresses. According to the estimate, a 1 percent increase in residential density leads to a 0.844 percent decrease of the conditional probability of migration.

Based on the structural relationship between the estimated elasticities, I further estimate the value of the structural parameter $\theta = -\beta_R/\beta_G$, which capture the extent of rivalry associated with local government goods and services, by the Delta method. The estimated value of parameter θ is 0.787 with a standard error equal to 0.315. The magnitude of the estimate suggests a non-negligible effect of rivalry from residential density.

Lastly, the estimated elasticity of worker mobility with respect to home prices is equal to -0.49, substatially larger than the OLS estimate in Column (1). As explained earlier, there

are two sources of bias to the OLS estimate of β_Q. One is the omitted variable bias. Because amenity values and home prices are positively correlated, the OLS estimate in Column (1) is biased upward towards zero. The other is measurement errors, attenuating the effect of home prices on worker mobility towards zero. The 2SLS estimates which correct for these issues show that the conditional probability of migration decreases by 0.49 percent as home prices increase by 1 percent.[32]

4.3   Interpretation of Estimates

The elasticity of worker's mobility to government expenditure has not been extensively estimated in the previous literature. Suárez-Serrato and Wingender (2014) estimate 1.46 for the elasticity of population at the county group level by leveraging exogenous variation in federal spending in the U.S. Fajgelbaum et al. (2019) obtains a similar value of the elasticity based on the number of workers at the state level in the U.S. Although the comparison is not perfect since they consider different source of variation, time periods, and geography, my estimate of 1.07 is close to their estimates.[33]

The literature on agglomeration economies includes population density as part of amenities and productivity (Ciccone and Hall, 1996; Glaeser and Maré, 2001; Ahlfeldt et al., 2015; de la Roca and Puga, 2017). The magnitude of its effect has been estimated to be positive, but rather small; the existing values of agglomeration parameter range from 0.01 to 0.06. This being said, the congestion parameter via local government goods θ in my model includes the agglomeration force.[34] Overall, I find that local government goods and services are rival. However, since it is not fully rival, a tax contribution from an additional resident is shared with all the other residents. Therefore, the effect of residential density on worker mobility is net agglomerating.

Lastly, the estimation results suggest that spatial frictions reflected in the iceberg costs of migration, commuting, and job finding are important determinants of worker's location decisions. In Table B.3 in Appendix, I report the OLS estimates including each of the fixed effects separately and show that the effects of local government expenditures, residential density, and home prices on worker mobility is sensitive to job finding costs in Column (2), migration costs in Column (3), and commuting costs in Column (4).

---

[32]In Table B.5 in Appendix, I report the estimation results based on an alternative, parsimonious specification in which time-varying origin-workplace fixed effects are replaced with fixed effects separately for time-by-origin, time-by-workplace, and origin-pair. The results are robust.

[33]My estimate of β_G is slightly smaller than the ones estimated in Suárez-Serrato and Wingender (2014) and Fajgelbaum et al. (2019). This is likely because they study the effects of government expenditures that affect firms as well as workers.

[34]There is a simple isomorphic formulation in which the agglomeration force is directly featured in the model. If the local amenities B_r is endogenous and depends on amenity fundamentals b_r and residential density R_r^γ, where γ captures the residential agglomeration force. Then, the reduced-form parameter $\beta_R = (\theta\lambda - \gamma)\epsilon$ and $\beta_R/\beta_G = \theta - \gamma/\lambda$.

In Table B.6 in Appendix, I compare the OLS and 2SLS estimates based on both migration and commuting flows in Column (1) and (2) to the OLS and 2SLS estimates based on migration flows alone in Column (3) and (4) and commuting flows alone in Column (5) and (6). The 2SLS estimates in Column (4) and Column (6) are biased because the exclusion restriction in each case is violated. On the one hand, based on the migration pattern alone, the effect of local government spending on the probability of migration is underestimated because workers move to places with higher commuting potentials, which compensate the lack of local government spending. On the other hand, the same effect is overestimated by about 5 times based on the commuting pattern alone because, in addition to direct cost of commuting, there exist migration and job finding costs that enable each commute. This set of estimation results emphasizes the importance of jointly accounting for both migration and commuting and for proper conditioning to consistently estimate the elasticities of worker mobility with respect to local government spending, residential density, and floor space prices.

5 Estimation of Spatial Frictions

Spatial frictions make it difficult for workers to reallocate across space. The model presented in Section 3 features the iceberg costs of worker mobility including three spatial frictions: the costs associated with migration, commuting, and job finding. In this section, I estimate the effects of spatial frictions on the spatial mobility of workers. I shed light on the importance of jointly considering migration and commuting decisions in correctly estimating the distance-elasticities of migration and commuting.[35]

5.1 Spatial Frictions in Migration and Commuting Decisions

I rewrite the gravity equation (5) by grouping the location-specific factors by residence ϕ_r, by workplace location ϕ_m, and by previous residence ϕ_o:

$$\pi_{orm} = \frac{\phi_o \phi_r \phi_m}{(\underbrace{\varepsilon_{orm} D_{or} D_{rm} D_{om}}_{D_{orm}})^\epsilon}. \tag{9}$$

I refer to Equation (9) as a generalized gravity equation of migration and commuting as this equation generalizes the gravity equations in the literature on migration and commuting.

Based on Equation (9), the expression for the spatial distribution of workers by their origins and current residences is given by:

---

[35]Recovering unobserved factors in the model requires the estimates of these elasticities as discussed in Section 7.3.

$$\pi_{or} = \frac{\phi_o \phi_r}{D_{or}^\epsilon} \underbrace{\sum_{m=1}^{J} \frac{\phi_m}{(\varepsilon_{orm} D_{rm} D_{om})^\epsilon}}_{ALMA_{or}\varepsilon_{or}} \cdot \qquad (10)$$

The key difference between the expression above (10) and the one considered in the literature on migration is the last term $\sum_{m=1}^{J} \phi_m/(\varepsilon_{orm} D_{rm} D_{om})^\epsilon$. This additional term can be expressed in terms of stochastic ε_{or} and systemic components. I refer the systemic component of the additional term to as *augmented labor market access* (*ALMA*). *ALMA* shares a similar structure with the labor market access (*LMA*) in Morten and Oliveira (2018) and more generally with the market access approach in Donaldson and Hornbeck (2016), but includes an additional factor D_{om}. On the one hand, the conventional *LMA* has a unique value for each of current residences (i.e., destinations) since it captures the benefit of accessing the local labor market net of commuting costs. On the other hand, *ALMA* allows *LMA* to vary by previous residences (i.e., origins) to account for heterogeneous costs of job finding and captures the benefit of accessing the local labor market net of both commuting and job finding costs.

ALMA captures the idea that workers from different origins value the same local labor market of a residence differently due to the cost of job finding. The extent to which workers can benefit from the labor market of a certain residence may depend on where they migrate from due to, for instance, a migrant network that makes job finding easier for workers from a certain origin relative to those from somewhere else (Card, 2001; Cadena and Kovak, 2016). Although *ALMA* does not explicitly appear in the gravity equations used in the migration literature, *ALMA* provides an important information about how workers sort across space. Workers conditional on their origins are more likely to migrate to a residence with higher *ALMA*, while a higher value of *ALMA* enables workers to afford a higher cost of migration.

Second, the literature on commuting employs a gravity equation, which summarizes the spatial distribution of workers in terms of their current residential and workplace locations. By summing π_{orm} in Equation (9) over initial residences, I obtain a gravity equation that characterizes the commuting patterns of workers:

$$\pi_{rm} = \frac{\phi_r \phi_m}{D_{rm}^\epsilon} \underbrace{\sum_{o=1}^{J} \frac{\phi_o}{(\varepsilon_{orm} D_{or} D_{om})^\epsilon}}_{AMMA_{rm}\varepsilon_{rm}} \cdot \qquad (11)$$

Again, the key difference between the gravity equation above (11) and the one considered in the literature on commuting is the last term $\sum_{o=1}^{J} \phi_o/(\varepsilon_{orm} D_{or} D_{om})^\epsilon$. This term can be written in terms of stochastic ε_{rm} and systemic components, last of which I term *augmented migrant (worker) market access* (*AMMA*). *AMMA* captures the average appeal of a commute (between

a residence and a workplace location) for migrants net of costs associated with migration and job finding. Therefore, there are two types of costs that explain the commuting patterns One type is a usual direct cost of commuting D_{rm}. The other is an indirect cost that captures the idea that it is costly to move to residence r and find a job in workplace location m from previous residence o in order to commute between r and m. Similar to the direct cost of commuting, this indirect cost makes the appeal of a commute less attractive.

$AMMA$ measures how accessible each commute is for workers originating from different places on average and varies at the commute-pair level.[36] On the one hand, it is likely to see more workers carrying out a certain commute when this commute has a higher value of $AMMA$. On the other hand, if the commute is costly, the appeal of this commute is lower and so is $AMMA$. While the literature on commuting is silent about the role of $AMMA$ as a determinant of commuting decisions, accounting for $AMMA$ is important to correctly estimate the distance elasticity of commuting.

5.2   Estimation Strategies and Results

I take a step towards evaluating how much spatial frictions quantitatively explain the spatial distribution of workers observed from the Population Census of South Korea. As defined in Section 3, I impose a structure on each of the bilateral linkages such that these linkages depend on distances d_{jk} between localities j and k as similarly done in, for instance, Morten and Oliveira (2018) for the cost of migration and Ahlfeldt et al. (2015) for the cost of commuting:

$$D_{or} = \exp(\rho d_{or}), \ \ D_{rm} = \exp(\kappa d_{rm}), \ \ D_{om} = \exp(\delta d_{om}). \tag{12}$$

The parameters ρ, κ, and δ control the sizes of migration, commuting, and job finding costs with respect to distances between spatial units. The motivation for imposing the same structure on the cost of job finding as the costs of migration and commuting is that finding a job is harder for workers who are located farther away from potential job sites. Taking into account that the data is available for cross-sections of 3 years (2005, 2010, and 2015), I augment the gravity equation by adding time subscripts:

$$\pi_{orm,t} = \frac{\phi_{r,t}\phi_{m,t}\phi_{o,t}}{\varepsilon_{orm,t}^{\epsilon}\exp(\rho\epsilon d_{or} + \kappa\epsilon d_{rm} + \delta\epsilon d_{om})}. \tag{13}$$

I estimate the reduced-form elasticities of worker mobility with respect to distances ($\rho\epsilon$, $\kappa\epsilon$, $\delta\epsilon$) using the South Korean Census. [37]

---

[36]The term *augmented migrant market access* reflects a concept that different commutes have differential capacities to access and attract migrants.

[37]In Appendix B.2, I conduct a type of decomposition exercise to shed light on the contribution of the spatial linkages jointly and discretely to the observed variation in the spatial distribution of workers. The spatial

5.2.1   Cost of Migration with respect to Distance

I take the log transformation of both sides of the generalized gravity equation (13)with time subscripts and the structure of bilateral linkages to obtain the expression as follows:

$$\ln \pi_{orm,t} = \phi_{rm,t} + \phi_{om,t} - \rho\epsilon d_{or} + \varepsilon^{mig}_{orm,t}, \tag{14}$$

where the current residence by workplace fixed effects interacted with year dummies $\phi_{rm,t}$ capture time-varying location specific factors at the current residence $\ln\phi_{r,t}$, the workplace $\ln\phi_{m,t}$, and the cost of commuting $-\kappa\epsilon d_{rm}$. The origin by workplace fixed effects interacted with year dummies $\phi_{om,t}$ capture time varying location specific factors at origin $\ln\phi_{o,t}$ as well as the cost of job finding $-\delta\epsilon d_{or}$. The parameter $\rho\epsilon$ is the semi-elasticity of migration flows with respect to distances of migration. The expected sign of $-\rho\epsilon$ is negative because workers are less likely to migrate to places that are farther away. The last term $\varepsilon^{mig}_{orm,t}$ corresponds to the log of the stochastic error $\varepsilon_{orm,t}$; I assume this error term is orthogonal to distances of migration.[38] I allow the errors to be correlated across migration pairs.

   In Table 4, I start with a simple OLS estimation without any fixed effects and gradually add two sets of fixed effects (commuting pairs $\phi_{rm,t}$ and job finding pairs $\phi_{om,t}$), one at a time. The estimate in Column (1) without any fixed effects is -0.002, statistically different from zero. This estimate is likely biased from omitting the determinants of migration that are correlated with distance of migration. For instance, if workers migrate longer distances to find better jobs (higher wages), the estimate is biased toward zero. In order to purge out the net benefits of living in r and commuting to workplace m, I include pairwise fixed effects for commuting pairs $\phi_{rm,t}$. The estimated coefficient is now slightly more negative at -0.004, reported in Column (2).

   In Column (3), I flexibly control for the cost of job finding by adding pairwise fixed effect for job finding pairs $\phi_{om,t}$; this specification corresponds to Equation (14). The estimated semi-elasticity is -0.033 and means that the probability of migration decreases by 3.3 percent with respect to a one-kilometer increase in the distance of migration. The large difference between the estimates in Column (2) and Column (3) implies that there exists a substantial upward bias rising from failing to account for the difficulty in finding jobs for workers who are migration from more distant places. The estimate in Column (3) captures the positive relationship between distance and the cost of migration, net of the costs associated with commuting and job finding.

---

linkages individually explain about 23 to 70 percent of the observed variation.

   [38]I estimate Equation (14) using a linear fixed effects estimator. The identification assumption is that, the distances of migration are uncorrelated with all other determinants of residential location choices conditional on the fixed effects. The error term may capture random measurement error in distances of migration. Although I do not observe exact distances of migration, the magnitude of potential measurement errors with respect to distance of migration are likely to be small because the geographical units are defined more finely compared to the spatial units considered in the previous literature.

Given that the distance of migration is a time-invariant feature that links the spatial units, I test whether or not the semi-elasticity of migration to distance varies over time. I include two additional regressors to Equation (14): distance interacted with dummy variables for year 2005 and 2010. The coefficients in front of the additional regressors tell us how different the semi-elasticities are in 2005 and 2010 relative to in 2015. The estimation result is reported in Column (4). The magnitudes of the estimated coefficients are economically small and statistically not different from zero. I conclude that the semi-elasticity of migration to distance is relatively constant, and therefore is a time-invariant feature describing the data.[39]

Lastly, I examine the consequence of using the probability of migration, a dependent variable commonly used in the previous literature on migration, to estimate the semi-elasticity of migration with respect to distance. I estimate a specification analogous to what the literature uses to estimate as follows:

$$\ln \pi_{or,t} = \tilde{\phi}_{r,t} + \tilde{\phi}_{ot} - \rho\epsilon d_{or} + \varepsilon_{or,t}^{mig}, \tag{15}$$

where the current residence and the origin fixed effects interacted with year dummies ($\tilde{\phi}_{r,t}$ and $\tilde{\phi}_{o,t}$) capture any push and pull factors specific to the origin and current residence that affect migration. To consistently estimate the semi-elasticity of migration to distance $-\rho\epsilon$, the error term $\varepsilon_{or,t}^{mig}$ must be orthogonal to either distance d_{or} or the dependent variable $\ln \pi_{or,t}$, or both. The gravity equation helps to unpack the error term. Based on Equation (10) with time subscripts on all the terms except distances, $\varepsilon_{or,t}^{mig}$ corresponds to $\ln ALMA_{or,t}\varepsilon_{or,t} = \ln \sum_{m=1}^{J} \frac{\phi_{m,t}}{(\varepsilon_{orm,t}D_{rm}D_{om})^\epsilon}$. An estimate without controlling for the effects of $ALMA_{or,t}$ on migration flows would be biased towards zero because, as explained above, $ALMA_{or,t}$ is correlated positively with both distance and the observed migration flows. I estimate Equation (14) and report the estimated coefficient in front of distance in Column (5). Conforming to the expected direction of the omitted variable bias, the estimate is only about a fifth of the estimate in Column (3) because workers are willing to migrate longer distances when they face higher returns from the local labor market at the destination.[40]

5.2.2   Cost of Commuting with respect to Distance

To estimate the semi-elasticity of commuting with respect to distance, I derive the following specification based on the generalized gravity equation:

---

[39]The results are robust to estimating the distance elasticity of migration pooling observations for each year (2005, 2010, and 2015).

[40]The estimate in Column (5) of Table 4 falls within the range of available estimates in the literature. Bryan and Morten (2018) estimates the elasticity of migration to distance in the U.S. (-0.553) and Indonesia (-0.717). Re-scaling the estimated semi-elasticity of -0.007 by the average migration distance (75.34 kilometers), the implied elasticity of migration to distance based on my estimate is -0.53.

$$\ln \pi_{orm,t} = \phi_{or,t} + \phi_{om,t} - \kappa\epsilon d_{rm} + \varepsilon_{orm,t}^{com}, \tag{16}$$

where the origin by current residence fixed effects interacted with year dummies $\phi_{or,t}$ capture time-varying location specific factors at the origin $\ln \phi_{o,t}$ and the current residence $\ln \phi_{r,t}$ as well as the cost of migration $-\rho\epsilon d_{or}$; the origin by workplace fixed effects interacted with year dummies $\phi_{om,t}$ capture time-varying location specific factors at the workplace $\ln \phi_{m,t}$ as well as the cost of job finding $-\delta\epsilon d_{om}$. The parameter $-\kappa\epsilon$ is the semi-elasticity of commuting flows with respect to distance of commuting. Because workers are less likely to commute longer distances from their location of residence, the sign of the semi-elasticity must be negative. The stochastic error term $\varepsilon_{orm,t}^{mig}$, orthogonal to distances of commuting, includes the log of the stochastic error. I allow the errors to be correlated across commuting pairs.

Table 5 report the estimation results. Like before, I start with a simple OLS estimation without any fixed effects and gradually add two sets of fixed effects (migration pairs $\phi_{or,t}$ and job finding pairs $\phi_{om,t}$), one at a time. The estimate without any fixed effects is -0.013, statistically different from zero in Column (1). This estimate is likely biased from omitting determinants of commuting flows that are correlated with distance of commuting. For example, workers who migrated from places farther away may not want to bear higher commuting costs in addition to cost of migration. Then, the estimate is biased toward zero. In order to account for the omitted variable bias associated with migration cost, I introduce the migration pair fixed effects in Column (2). As expected, the estimate reported in Column (2) is -0.035, more negative compared to the estimate in Column (1). Furthermore, the returns from working in m net of job finding cost, captured by $D_{om,t}$ are positively correlated with the commuting flows and allows workers to afford higher commuting cost. This implies another bias toward zero.

In order to address this issue, Column (3) estimates the semi-elasticity of commuting flows with respect to distance with both fixed effects of migration and job finding pairs. The estimated elasticity in Column (3) is -0.045: a one-kilometer increase in commuting distance decreases the probability of commuting by 4.5 percent.[41] To understand how stable the semi-elasticity of commuting to distance over time is, I additionally include distance interacted with year dummy variables for 2005 and 2010. The estimation results in Column (4) indicate that the semi-elasticity of commuting with respect to distance is stable over time.

Next, I examine what happens if the bilateral linkages of migration and commuting are not accounted for when estimating the semi-elasticity of commuting with respect to distance. To

---

[41]Travel time is also widely used to define a cost of geographical mobility (Ahlfeldt et al., 2015; Morten and Oliveira, 2018). In Appendix B.4, I show that travel time associated with commuting has a one-to-one relationship with distance of commuting. I re-estimate Equation (16) by using commute time reported in the Population census as an endogenous regressor, instrumented with distance of commuting. The estimated semi-elasticity (-0.036) is stastistically not different from the estimate in Column (3) of Table 5 (-0.033).

do so, I follow the literature on commuting and use the probability of commuting $\ln \pi_{rm,t}$ as a dependent variable and estimate the following specification:

$$\ln \pi_{rm,t} = \tilde{\phi}_{r,t} + \tilde{\phi}_{m,t} - \kappa\epsilon d_{rm} + \varepsilon_{rm,t}^{com}, \tag{17}$$

where the residence and the workplace fixed effects interacted with year dummies ($\tilde{\phi}_{r,t}$ and $\tilde{\phi}_{m,t}$) capture any factors specific to residence and workplace that affect commuting (costs of living and wages). In order to consistently estimate the semi-elasticity of commuting to distance $\kappa\epsilon$, the error term $\varepsilon_{rm,t}^{com}$ must be uncorrelated to either distance d_{rm} or the probability of commuting $\ln \pi_{rm,t}$, or both. Similar to the case of migration, the log of Equation (11) with time subscripts whenever applicable has a direct correspondence with Equation (17). The residual term $\varepsilon_{rm,t}^{com}$ is equal to $\ln AMMA_{rm,t}\varepsilon_{rm,t} = \ln \sum_{o=1}^{J} \frac{\phi_{o,t}}{(\varepsilon_{orm,t}D_{or}D_{om})^\epsilon}$. It is clear that an increase in $\ln AMMA_{rm,t}$ increases the probability of commuting. Estimating $\kappa\epsilon$ without controlling for the effects of $\ln AMMA_{rm,t}$ on commuting flows would be biased away from zero if a high commuting cost is associated with a low value of $AMMA_{rm,t}$. Column (5) reports the estimated semi-elasticity based on Equation (17). The estimate is -0.074, which is more negative compared to the estimate in Column (3) in line with the intuition.[42]

5.2.3 Cost of Job Finding with respect to Distance

In this subsection, I estimate the semi-elasticity of job finding with respect to distance. I derive an estimating equation by taking the log transformation of Equation (13):

$$\ln \pi_{orm,t} = \phi_{rm,t} + \phi_{or,t} - \delta\epsilon d_{om} + \varepsilon_{orm,t}^{jf}, \tag{18}$$

where the commute-pair fixed effects $\phi_{rm,t}$ capture net benefits of living in r and working in m, $\ln \frac{\phi_{r,t}\phi_{m,t}}{D_{rm}^\epsilon}$ (e.g., housing prices, wages, and commuting cost); the migration-pair fixed effects $\phi_{or,t}$ capture the cost of migration $-\rho\epsilon d_{or}$ as well as any factors that make o a more or less attractive residence to stay $\ln \phi_{o,t}$; the sign of the parameter $\delta\epsilon$ is likely positive because it is harder to find jobs that are farther away from where workers migrate; the last term $\varepsilon_{orm,t}^{jf}$ captures the random noise. I allow the errors to be correlated across job finding pairs.

The simple OLS estimate of the semi-elasticity of job finding with respect to without any fixed effects is -0.001, reported in Column (1) of Table 6. Workers are willing to accept a high cost of job finding (equivalently, a large d_{om}) if doing so allows them to find a pair of residence and workplace locations with higher wage, lower cost of living and lower commuting

---

[42]The estimate in Column (5) of Table 5 is close to the available estimates in the literature. Ahlfeldt et al. (2015) estimates the same semi-elasticity based the inter-district commuting flows in Berlin, Germany in 2008 contemporaneous to the time period considered in this paper. Their estimated semi-elasticity of commuting with respect to distance (also measured in kilometer) is equal to -0.07.

costs. These correlations results in bias towards zero. In Column (2), I introduce the pairs fixed effects for residence and workplace locations and find an estimate more negative, compared to Column (1). Furthermore, because the distance of migration and the distance of job finding are positively correlated and workers are less like to migrate farther away, the estimate in Column (2) is still biased toward zero.

Column (3) reports the estimated semi-elasticity of job finding with respect to distance with both fixed effects as prescribed in Equation (18). According to Column (3), the probability of job finding decreases by 1.6 percent for a one-kilometer increase in distance of job finding. To examine how stable the semi-elastisticity of job finding is with respect to distance, I introduce distances interacted with year dummy variables for 2005 and 2010. In Column (4), the coefficient estimate for distance reported in the first row is the semi-elasticity of job finding to distance in 2015. The difference between the estimate in Column (4) and the estimate reported in Column (3) is economically small and is not statistically significant.[43]

5.3   Implications

Taking the gravity framework to the spatial distribution of workers in South Korea, I find that the spatial linkages between localities (costs of migration, commuting, and job finding) are important determinants of the spatial distribution of workers. In particular, they are systematically explained by distances between spatial units. The estimated reduced-form elasticities are negative and stable over time.[44]

I make two important distinctions from the previous literature on migration and commuting. First, residence does not need to be a place for both living and working. Second, where workers come from matters for not only determining where they live, but also where they work today.[45] I find substantial biases with the estimates of the distance elasticities of migration and commuting reported in the previous literature.

First, the estimated elasticities of migration available in the literature are likely biased

---

[43]To the best of my knowledge, there is no existing estimate of the decay parameter $\delta\epsilon$ (i.e. elasticity of job finding with respect to distance) in the literature. That being said, my estimate of the spatial decay of job finding can be considered as a reduced-form parameter combining the effects of distance on job match (employment) and job application (intent for employment), last of which Manning and Petrongolo (2017) estimate based on a spatial model of job search using the data on the demand and supply of the job search process in the U.K. They find a relatively strong decay of job applications in distance. Marinescu and Rathelot (2018) also finds similar results (implied semi-elasticity of job application to distance equal to 0.02) in the context of the U.S.

[44]This finding, in particular related to migration, is consistent with the assumption on migration friction in Caliendo et al. (2019), which build the sectoral mobility costs in Dix-Carneiro (2014). Ahlfeldt et al. (2015) make the same assumption about the semi-elasicity of commuting and applies the same spatial decay of commuting estimated based on the commuting patterns of workers in the city of Berlin in 2008 to explain the commuting patterns before and after the division and reunification of East and West Germany.

[45]Also, Pellegrina and Sotelo (2019) find that the origins of agricultural workers matter in determining the types of crops they cultivate they migrate to a different region in the context of Brazil.

toward zero because the cost of migration is positively correlated with the benefits from changing residences discounted by costs associated with commuting and job finding ($ALMA$). This means that workers appear to be willing to migrate longer distance for better labor market access. Second, the available estimates for distance elasticity of commuting in the literature are likely biased away from zero (more negative). Because of omitting the appeal of commuting net of indirect costs rising from migration and job finding that enable a certain commute ($AMMA$), workers appear to be more sensitive to commuting distance than they actually are.[46]

6 Quantitative Spatial General Equilibrium Model

I take a step towards quantifying the welfare consequences of the fiscal arrangements observed in 2015. Accordingly, I embed the partial equilibrium model of worker's location decisions presented in Section 3 into a general equilibrium setup. I model the production of consumption goods and the allocation of floor spaces for residential and commercial use. Local government spending is determined based on national policies on taxation, revenue sharing, and the rules of redistribution. In equilibrium, wages, floor space prices, and local government expenditures are endogenously determined along with the spatial distribution of workers. Lastly, I define the spatial general equilibrium of the economy.

6.1 More on Worker's Location Decisions

I characterize the market clearing conditions for migration and commuting based on the gravity equation (5) derived in Section 3. First, summing the probabilities of choosing residence r and workplace m conditional on moving from origin o across workplaces, I obtain the expression for the probabilities of moving to r given origin o:

---

[46]The results also shed light on timings of mobility decisions. Intuitively, there are two alternative timings of how workers decide where to live and where to work. First, a worker may decide a residence where he would like to live (including the option to stay), and then find a job. If this timing is true, the semi-elasticity of job finding should be estimated to zero controlling for the commuting-pair fixed effects. Second, a worker may find a job first, then decide where to commute from. If this alternative timing is true, then the semi-elasticity of commuting should be estimated similarly with or without the fixed effects accounting for the job finding cost conditional on the migration pair fixed effects. Both of these alternative timings are inconsistent with the observed spatial distribution of workers. The findings altogether imply that a certain timing assumption is too restrictive to explain the variations in the observed spatial distribution of workers. Consistent with these findings, the model presented in this paper allows workers to make migration and commuting decisions jointly.

$$\pi_{r|o} = \sum_{m=1}^{J} \frac{\pi_{orm}}{\pi_o} = \frac{\sum_{m=1}^{J} \Phi_{orm}}{\Phi_o}$$

$$= \frac{T_r \left(\frac{B_r}{D_{or} Q_r^{1-\beta}} \left(\frac{G_r}{R_r^{\theta}} \right)^{\lambda} \right)^{\epsilon} \sum_{m=1}^{J} M_m \left(\frac{(1-\tau_m) w_m}{D_{om} D_{rm}} \right)^{\epsilon}}{\sum_{r'=1}^{J} T_{r'} \left(\frac{B_{r'}}{D_{or'} Q_{r'}^{1-\beta}} \left(\frac{G_{r'}}{R_{r'}^{\theta}} \right)^{\lambda} \right)^{\epsilon} \underbrace{\sum_{m'=1}^{J} M_{m'} \left(\frac{(1-\tau_{m'}) w_{m'}}{D_{r'm'} D_{om}} \right)^{\epsilon}}_{ALMA_{or'}}}.$$

Workers are more like to migrate a residence with a higher amenity value B_r, a higher benefit from local government goods $\frac{G_r}{R_r^{\theta}}$, and a lower per-unit price of floor space Q_r. In addition, there are two sources of bilateral determinants. The probability of choosing residence r decreases in the cost of migration D_{or}, but increases in the benefit of accessing the labor market discounted by commuting and job finding costs $\sum_{m=1}^{J} M_m \left(\frac{(1-\tau_m) w_m}{D_{om} D_{rm}} \right)^{\epsilon}$, which corresponds to the augmented labor market access $ALMA_{or}$. Using these conditional probabilities, migration market clearing condition requires that the number of workers who live in r is equal to the sum of workers migrating to r from all possible origins o:

$$R_r = \sum_{o=1}^{J} \pi_{r|o} R_o = \sum_{o=1}^{J} \frac{\left(\frac{\tilde{B}_r}{D_{or} Q_r^{1-\beta}} \left(\frac{G_r}{R_r^{\theta}} \right)^{\lambda} \right)^{\epsilon} ALMA_{or}}{\sum_{r'=1}^{J} \left(\frac{\tilde{B}_{r'}}{D_{or'} Q_{r'}^{1-\beta}} \left(\frac{G_{r'}}{R_{r'}^{\theta}} \right)^{\lambda} \right)^{\epsilon} ALMA_{or'}} R_o. \qquad (19)$$

I derive the expression for the probability of commuting commuting to workplace m conditional on living in residence r. I take the ratio of the unconditional joint distribution of workers in terms of their residence and workplace to the the unconditional distribution of workers by residence as follows:

$$\pi_{m|r} = \frac{\sum_{o=1}^{J} \pi_{orm}}{\sum_{m'=1}^{J} \sum_{o'=1}^{J} \pi_{o'r'm'}} = \frac{\sum_{r_0=1}^{J} \Phi_{orm} \pi_o / \Phi_o}{\sum_{m'=1}^{J} \sum_{o'=1}^{J} \Phi_{o'rm'} \pi_{o'} / \Phi_{o'}}$$

$$= \frac{\left(\frac{(1-\tau_m) \tilde{w}_m}{D_{rm}} \right)^{\epsilon} \sum_{o=1}^{J} \frac{\pi_o / \Phi_o}{(D_{or} D_{om})^{\epsilon}}}{\sum_{m'=1}^{J} \left(\frac{(1-\tau_{m'}) \tilde{w}_{m'}}{D_{rm'}} \right)^{\epsilon} \underbrace{\sum_{o'=1}^{J} \frac{\pi_{o'} / \Phi_{o'}}{(D_{o'r} D_{o'm})^{\epsilon}}}_{AMMA_{rm'}}},$$

where the terms specific to current residence such as amenities, housing prices, and government goods are canceled out from the numerator and denominator. In line with intuition, workers are more likely to commute to places with higher returns $((1-\tau_m) \tilde{w}_m)^{\epsilon}$ net of commuting costs

31

D_{rm}. Moreover, the conditional probability of commuting depends on how costly it is to migrate to residence r and find a job in workplace m, $\sum_{o=1}^{J} \frac{\pi_o / \Phi_o}{(D_{or} D_{om})^{\epsilon}}$, which corresponds to augmented migrant market access $AMMA$. Using these probabilities, I obtain the following expression:

$$L_m = \sum_{r=1}^{J} \pi_{m|r} R_r = \sum_{r=1}^{J} \frac{\left(\frac{(1-\tau_m) \tilde{w}_m}{D_{rm}} \right)^{\epsilon} AMMA_{rm}}{\sum_{m'=1}^{J} \left(\frac{(1-\tau_{m'}) \tilde{w}_{m'}}{D_{rm'}} \right)^{\epsilon} AMMA_{rm'}} R_r, \tag{20}$$

where the number of workers employed in m is equated with the number of workers choosing to commute to m from all possible residences. I refer to this equation as the commuting market clearing condition.

Expected income of workers living in district r is equal to the sum of the after-tax wages in all possible workplace locations weighted by the conditional probabilities of commuting to those locations:

$$\mathbb{E}[(1-\tau_m) w_m | r] = \sum_{m=1}^{J} \frac{\left(\frac{(1-\tau_m) \tilde{w}_m}{D_{rm}} \right)^{\epsilon} AMMA_{rm}}{\sum_{m'=1}^{J} \left(\frac{(1-\tau_{m'}) \tilde{w}_{m'}}{D_{rm'}} \right)^{\epsilon} AMMA_{rm'}} (1-\tau_m) w_m. \tag{21}$$

Expected income of workers are higher in places with lower costs of commuting D_{rm} as well as higher $AMMA_{rm}$, the indirect cost of commuting rising from the costs associated with migration and job finding. Because workers allocate $1 - \beta$ fraction of their income to housing, the demand for residential floor space is given by

$$H_r^R = (1-\beta) \frac{\mathbb{E}[(1-\tau_m) w_m | r] R_r}{Q_r}. \tag{22}$$

Lastly, the population mobility implies that the ex-ante expected utility for each initial residence is the same across all possible residence-workplace pairs. That is,

$$\mathbb{E}[u_o] = \Gamma\left(\frac{\epsilon-1}{\epsilon}\right) \Phi_o^{1/\epsilon} = \Gamma\left(\frac{\epsilon-1}{\epsilon}\right) \left[\sum_{r'=1}^{J} \sum_{m'=1}^{J} \left(\frac{\tilde{B}_{r'} (1-\tau_{m'}) \tilde{w}_{m'}}{D_{or'm'} Q_{r'}^{1-\beta}} \left(\frac{G_{r'}}{R_{r'}^{\theta}} \right)^{\lambda} \right)^{\epsilon} \right]^{1/\epsilon} \equiv \bar{u}_o, \tag{23}$$

where the expectation is taken over the distribution of the idiosyncratic component of utility.[47] I construct a measure of economy-wide welfare by taking the average of the expected utilities (23) weighted by the distribution of workers by their origins π_o: $\bar{u} = \sum_{o=1}^{J} \bar{u}_o \pi_o$. This measure corresponds to consumption equivalent worker welfare.

---

[47]See Appendix D.1 for the derivation of Equation (23).

6.2 Production

The production of the tradable final good occurs under conditions of perfect competition and constant returns to scale. In particular, I assume that the production technology follows Cobb-Douglas as follows:

$$y_m = A_m L_m^\alpha H_m^{F\,1-\alpha} \tag{24}$$

where A_m is final goods productivity; L_m is labor input; and H_m^F corresponds to a measure of floor space used commercially. Profit maximization under perfect competition implies that labor demand is high in places where productivity A_m is high; and wages w_m are lower in places with higher floor space available for commercial use H_m^F. This is captured in the labor demand as follows:

$$L_m = \left(\frac{\alpha A_m}{w_m}\right)^{\frac{1}{1-\alpha}} H_m^F. \tag{25}$$

The equilibrium wage equates the labor demand (25) to the labor supply (20) in each location. Similarly, the demand for floor space is given by

$$H_m^F = \left(\frac{(1-\alpha)A_m}{Q_m}\right)^{\frac{1}{\alpha}} L_m. \tag{26}$$

The demand for floor space is high in a district with the low equilibrium floor space price Q_m, high productivity A_m, and measure of workers L_m.

6.3 Floor Space Market Clearing

There is a fixed floor space for each district H_j, which can be used residentially and commercially. Atomistic absentee landlords allocate ϑ_j fraction of H_j to commercial use and $1-\vartheta_j$ to residential use. Therefore, market clearing for residential floor space requires that the demand and supply of residential space are equal to each other (i.e., $H_j^F = (1-\vartheta_j)H_j$):

$$(1-\beta)\frac{\mathbb{E}[(1-\tau_m)w_m|r]R_j}{Q_j} = (1-\vartheta_j)H_j. \tag{27}$$

Commercial floor space market clearing requires that the demand for commercial floor space equals the supply of floor space allocated to commercial use (i.e., $H_j^R = \vartheta_j H_j$):

$$\left(\frac{(1-\alpha)A_j}{Q_j}\right)^{\frac{1}{1-\alpha}} L_j = \vartheta_j H_j. \tag{28}$$

The setup of the floor space market in my model is consistent with the standard approach in the urban literature of assuming fixed supply (Rosen, 1979; Roback, 1982; Tsivanidis, 2019) and

33

allowing residential and commercial uses (Ahlfeldt et al., 2015; Monte et al., 2018; Tsivanidis, 2019).[48]

6.4 National and Local Governments

Consistent with the national fiscal policies discussed in 2.3, I model how local government expenditures are determined. First, the national government determines a progressive income tax schedule $\tau(w)$, which is increasing in w, for all districts to levy their residents and collects the fraction of local tax revenue $1 - \varsigma$ from each district. This means that ς fraction of total local tax revenue is kept locally, while $1 - \varsigma$ fraction is delivered to the national government. I refer the parameter ς to as local-national revenue sharing. Also, without loss of generality, I express $\tau(w_m) = \tau_m$.

Second, the national government operates intergovernmental transfers to supplement tax revenues retained locally. It allocates χ fraction of the national tax revenue (or equivalently, $(1 - \varsigma)\chi$ fraction of total local tax revenue) for redistribution via intergovernmental transfers. Then, the national government determines the shares ς_j of the budget allotted for intergovernmental transfers to be delivered to each local governments such that $\varsigma_j \geq 0$ for all $j = 1, ..., J$ and $\sum_{j=1}^{S} \varsigma_j = 1$. I refer to $\{\varsigma_j\}_{j=1}^{J}$ as rules of redistribution. Lastly, the national government uses $(1 - \varsigma)(1 - \chi)$ fraction of total local tax revenue to provision national government goods and services such as national defense and diplomacy. I assume that national government goods and services benefit workers equally regardless of where workers live and work.

Given the national fiscal policies $\{\{\tau_m\}_{m=1}^{J}, \varsigma, \chi, \{\varsigma_j\}_{j=1}^{J}\}$, measure of workers living in j (R_j), conditional probabilities of commuting ($\{\pi_{m|j}\}_{m=1}^{J}$), and wages ($\{w_m\}_{m=1}^{J}$) determine local government budget in district j. A budget balancing equation of local government in district j is expressed as follows

$$G_j = \varsigma \underbrace{\sum_{m=1}^{J} \tau_m w_m \pi_{m|j} R_j}_{TR_j} + \varsigma_j (1 - \varsigma)\chi \sum_{j'=1}^{J} TR_{j'}, \tag{29}$$

where $\sum_{m=1}^{J} \tau_m w_m \pi_{m|j} R_j$ is equal to local tax revenue collected from workers living in district

---

[48]The choice to assume a fixed stock of floor space for each district is to focus on evaluating the consequence of spatial distribution of government spending. It is reasonable to assume that the total stock of floor space does not adjust instantly. While the total stock for each district is fixed, the model allows its allocation to residential and commercial uses to vary. As discussed in Redding and Rossi-Hansberg (2017), assuming absentee landlord following the urban economics literature does not allow the model to capture full general equilibrium effects. In addition, in my model, a single floor space price for each unit clears the floor space market clearing conditions for both the residential and commercial floor space markets. An extension to the model can be easily made to incorporate land use regulations limit the return to floor space allotted to commercial use as in Ahlfeldt et al. (2015) and Tsivanidis (2019).

j denoted by TR_j. Therefore, the first term corresponds to local tax revenue collected and retained by local government in district j. The second term is the amount of intergovernmental transfers from the national government, equal to the redistribution parameter for district j (ς_j) multiplied by the total budget allotted for intergovernmental transfers, $(1-\varsigma)\chi\sum_{j'=1}^{S}TR_{j'}$. The extent of fiscal decentralization is captured by $\tilde{\chi} = \varsigma + (1-\varsigma)\chi$, which corresponds to the fraction of total tax revenue spent locally.

Depending on the rules of redistribution, local government expenditure in a district may be greater if $\varsigma_j > TR_j/\sum_{j'=1}^{J}TR_{j'}$ or less than its contribution to intergovernmental transfers. In this sense, the spatial distribution of local government spending is considered as a consequence of transfers across districts. The redistribution mechanism described in this section has features that are structurally similar to a transfer scheme based on lump-sum tax and government spending laid out in Fajgelbaum and Gaubert (2018) and more broadly place-based policies (Glaeser and Gottlieb, 2008; Kline and Moretti, 2014).

6.5 General Equilibrium

Given vectors of exogenous location characteristics $\{T_j, M_j, B_j, A_j, d_{jk}, H_j\}$, initial distribution of workers $\{\pi_o\}$, total measure of workers R, national fiscal policies $\{\tau_j, \varsigma, \chi, \varsigma_j\}$, and model parameters $\{\alpha, \beta, \lambda, \theta, \kappa, \rho, \delta, \epsilon\}$, a general equilibrium of this economy is defined as a vector of endogenous objects $\{R_j, L_j, w_j, Q_j, \vartheta_j, G_j, \bar{u}_o\}$. These seven components of the equilibrium vector are determined by the migration market clearing (19), commuting market clearing (20), labor market clearing (25), floor space market clearing for residential and commercial uses (27 and 28), local government budget balancing equation (29), and population mobility (23).

7 Parameterization of the GE Model

So far, I have estimated one structural parameter governing the extent of rivalry associated with benefits from local government spending (θ) in Section 4.2 and five reduced-form elasticities: the elasticities of worker mobility to local government expenditure ($\lambda\epsilon$) and to home prices ($(1-\beta)\epsilon$) in Section 4.2 and the semi-elasticities of migration, commuting, and job finding with respect to distance ($\rho\epsilon$, $\kappa\epsilon$, $\delta\epsilon$) in Section 5.2. In this section, I discuss how I estimate the rest of the model parameters and recover unobserved local characteristics for year 2015.

7.1 Labor Share in Production and Housing Expenditure Share

First, the labor share (α=0.823) is estimated by computing average share of labor cost to the total costs across districts reported in Economic Census in 2015, consistent with the findings of

Ákao Valentinyi and Herrendorf (2008). Second, I set housing expenditure $1 - \beta$ equal to 0.15 to match the observed housing expenditure share based on Household Expenditure Survey in 2015. This value is corroborated with the reported value reported in OECD (2016).

7.2 National Fiscal Policy Parameters

The values of the national policy parameters are directly observed in a collection of laws governing local fiscal capacities (the Local Tax Act and the Local Subsidy Act). In 2015, $\varsigma = 9.1\%$ of local tax revenue retained after tax collection according to the Local Tax Act. The Local Subsidy Act allocates $\chi = 35\%$ of total local tax revenue delivered to the national government for redistribution. Because I observe the amount of intergovernmental transfers (IT_j) for each district, I recover the values for the redistribution parameters as follows:

$$\varsigma_j = \frac{IT_j}{\sum_{j'=1}^{J} IT_{j'}}. \tag{30}$$

The last national policy parameter of interest is the tax rates. The tax rates by income brackets are observed in the Income Tax Act as discussed in Section 2.3. However, the observed tax rates cannot be directly used because I do not observe the distribution of wages within each district, nor does the model feature wage dispersion within each locality. Without relying on the observed tax rates, I solve for tax rates τ_m by district based on the observed local tax revenue (LT_r), probability of commuting ($\pi_{m|r}$), wages (w_m), and number of workers by residence (R_r) by inverting the following system of equations:

$$\frac{1}{\varsigma}\begin{bmatrix} LT_{r=1} \\ \vdots \\ LT_{r=J} \end{bmatrix} = \left(\begin{bmatrix} \tau_{m=1} \\ \vdots \\ \tau_{m=J} \end{bmatrix} l_{1\times J} I_{J\times J} \begin{bmatrix} w_{m=1} \\ \vdots \\ w_{m=J} \end{bmatrix}\right)' \begin{bmatrix} \pi_{m=1|r=1} & \cdots & \pi_{m=J|r=1} \\ \vdots & \ddots & \vdots \\ \pi_{m=1|r=J} & \cdots & \pi_{m=J|r=J} \end{bmatrix} \begin{bmatrix} R_{r=1} \\ \vdots \\ R_{r=J} \end{bmatrix}, \tag{31}$$

where $l_{1\times J}$ is a vector with all of its elements equal to 1. Finally, I simplify the tax rates to a single index ($\tau_m = \tau \ \forall m$) and calibrate the simplified tax rate equal to 0.245, the average tax rates from the inversion weighted by number of workers.[49]

---

[49]The counterfactual policy experiments concern with changes in the spatial distribution of spending due to changes in the intensity of fiscal decentralization and redistribution whiling holding the nationally determined tax rates fixed.

7.3 Recovery of Unobserved Local Characteristics

7.3.1 Local Productivity

I recover the values for local productivity using the observed wages and floor space prices. To satisfy the profit maximization and zero profit conditions, equilibrium floor space prices must satisfy:

$$Q_j = (1 - \alpha) \left(\frac{\alpha}{w_j} \right)^{\frac{\alpha}{1-\alpha}} A_j^{\frac{1}{1-\alpha}}. \tag{32}$$

Therefore, given the observed data on wages and floor space prices in 2015 and the parameter value of α, I can recover A_j for each district using the equilibrium condition above (32). Figure C.1 in Appendix C.2 plots the spatial distribution of the recovered values of local productivity. The greater Seoul area, the Northwestern part of South Korea, has relatively greater values of productivity, as well as the some of the coastal districts with ports (e.g., the greater Busan area covering the Southeastern coast) consistent with coastal and port advantages studied in Balboni (2019) and Ducruet et al. (2019).

7.3.2 Fréchet Shape Parameter

I estimate the Fréchet shape parameter, which is equivalent to the elasticity of worker mobility with respect to wage. I begin by deriving the expression for the probabilities of working in m conditional on living in r and having moved from o:

$$\pi_{m|ro} = \frac{\pi_{orm}}{\sum_{m'=1}^{J} \pi_{orm'}} = \frac{\Phi_{orm}}{\sum_{m'=1}^{J} \Phi_{orm'}} = \frac{\frac{\tilde{w}_m^\epsilon}{\exp(\kappa \epsilon d_{rm} + \delta \epsilon d_{om})}}{\sum_{m'=1}^{J} \frac{\tilde{w}_{m'}^\epsilon}{\exp(\kappa \epsilon d_{rm'} + \delta \epsilon d_{om'})}}. \tag{33}$$

I define a composite referred to as adjusted wages $\omega_j = \tilde{w}_j^\epsilon = M_j w_m^\epsilon$. I rewrite the above equation using adjusted wages and take the log transformation of both sides. Using my estimates of $\kappa \epsilon = 0.045$ and $\delta \epsilon = 0.016$ and rearranging such that left hand side consists of only observables, I obtain the following expression:

$$\ln \pi_{m|ro} + \kappa \epsilon d_{rm} + \delta \epsilon d_{om} = -\ln \sum_{m'=1}^{J} \frac{\omega_{m'}}{\exp(\kappa \epsilon d_{rm'} + \delta \epsilon d_{om'})} + \ln \omega_m, \tag{34}$$

where I treat $\kappa \epsilon d_{rm} + \delta \epsilon d_{om}$ as data and I observe $\ln \pi_{m|ro}$. The left hand side altogether can be decomposed into two parts: the first term that varies at the current residence and origin level and the second term that varies at the workplace level. Introducing stochastic errors to Equation (34), I regress the left hand side on the pairwise fixed effects of current residence and origin and the workplace fixed effects. Then, I recover the values of log adjusted wages from

the estimated workplace fixed effects. Note that these values are determined independent of ϵ based on the observed distribution of workers and the costs of commuting and job finding.[50]

The parameter ϵ controls the variance of log adjusted wages $(\ln \omega_j)$ relative to the variance of log observed wages $(\ln w_m^\epsilon)$. That is, $\sigma^2_{\ln w_j} = \frac{1}{\epsilon^2}\sigma^2_{\ln \omega_j}$ because the parameters M_j are deterministic. Therefore, I estimate the value of ϵ by taking the ratio of the standard deviations of log adjusted wages and log wages in the data after normalizing both to have geometric mean equal to 1. The resulting value of ϵ is equal to 3.54; this means the worker mobility increases by 3.54 percent for a 1 percent increase in wages.[51]

There are several other papers which estimate the same parameter. Defining spatial units as U.S. counties from 2006 to 2010, Monte et al. (2018) finds a point estimate of the shape parameter equal to 3.3, while Ahlfeldt et al. (2015) estimate its value equal to 6.83 based on the inter-district commuting patterns in the city of Berlin in 2008. My estimate falls within the responsible range of the existing estimates in the literature. With the estimated value of ϵ, I recover the structural parameters $(\lambda, \rho, \kappa, \rho)$ from the estimated reduced-form elasticities $(\lambda = \tilde{\lambda}/\epsilon = 0.30, \ \rho = \tilde{\rho}/\epsilon = 0.009, \ \kappa = \tilde{\kappa}/\epsilon = 0.013, \ \delta = \tilde{\delta}/\epsilon = 0.005)$. Furthermore, the estimated elasticity of worker mobility with respect to floor space prices reported in Table 3 is equal to $(1 - \beta)\epsilon$. Based on the estimate of $\epsilon = 3.54$, the implied value of $1 - \beta$ is equal to 0.14, close to the expenditure share estimated using the Household Expenditure Survey in Section 7.1.

Based on the structural value of how much people of local government spending $(\lambda = 0.3)$, I obtain the valuation of local government spending by computing the compensating variation. At the median values of per-capita local government spending (7,302 USD) and household income (18,180 USD) in 2015, workers are willing to give up 75 cent for a dollar increase in per-capita local government expenditure in their residence.

7.3.3   Adjusted Local Amenities

I recover adjusted amenity for each residence that rationalizes the observed spatial distribution of workers. Similarly to the process described when recovering the adjusted wages, I begin by deriving the expression for the conditional distribution of workers by their residences on workplace location and previous residence based on the gravity equation (5):

---

[50]In Appendix C.2, I examine the relationship between the recovered values of local productivity and number of firms in Figure C.2. Panel (a) shows that districts with higher productivity values have higher number of firms. In Panel (b), productivity is positively correlated with number of firms which discharge wastewater.

[51]For inference, I randomly sample 111 observations from log observed wages and log adjusted wages and compute ϵ. I repeat this process a number of times, e.g., 1 billion times, to obtain the distribution of the estimator for ϵ. Given the estimated standard error equal to 0.102. I reject the null hypothesis that $\epsilon = 0$ at the 99% confidence level. The result is robust to different sample sizes (25, 50, 100, 150 and 200).

$$\pi_{r|mo} = \frac{\pi_{orm}}{\sum_{r'=1}^{J} \pi_{or'm}} = \frac{\Phi_{orm}}{\sum_{r'=1}^{J} \Phi_{or'm}} = \frac{\frac{\tilde{B}_r^\epsilon G_r^{\lambda\epsilon}}{\exp(\kappa\epsilon d_{rm} + \rho\epsilon d_{or})Q_r^{(1-\beta)\epsilon} R_r^{\theta\lambda\epsilon}}}{\sum_{r'=1}^{J} \frac{\tilde{B}_{r'}^\epsilon G_{r'}^{\lambda\epsilon}}{\exp(\kappa\epsilon d_{r'm} + \rho\epsilon d_{or'})Q_{r'}^{(1-\beta)\epsilon} R_{r'}^{\theta\lambda\epsilon}}}. \tag{35}$$

I take the log transformation of both sides of Equation (35) and rearrange such that left hand side only consists of observables:

$$\ln \pi_{r|mo} - \ln \frac{G_r^{\lambda\epsilon}}{\exp(\kappa\epsilon d_{rm} + \rho\epsilon d_{or})Q_r^{(1-\beta)\epsilon} R_r^{\theta\lambda\epsilon}} = -\ln \sum_{r'=1}^{J} \frac{\tilde{B}_{r'}^\epsilon G_{r'}^{\lambda\epsilon}}{\exp(\kappa\epsilon d_{r'm} + \rho\epsilon d_{or'})Q_{r'}^{(1-\beta)\epsilon} R_{r'}^{\theta\lambda\epsilon}} + \ln \tilde{B}_r^\epsilon, \tag{36}$$

where I treat the second term in the left hand side as data given the parameter values. Introducing stochastic errors, I regress the left hand side on the pairwise fixed effects of workplace and origin and the residence fixed effects. Then, I recover the values of log adjusted amenities from the estimated residence fixed effects (up to scale). Figure C.3 in Appendix C.3 plots the spatial distribution of adjusted amenities. The metropolitan areas tend to have relatively higher amenity values, reflecting urban amenities. Also, the amenities are higher in the coastal areas, especially the coastal districts in the East and South.[52]

7.4 Non-targeted Moments

I evaluate how well the model predicts the non-targeted moments. First, I compare the observed data on number of workers by employment location to the model prediction in Panel (a) and (b) of Figure 5. The two variables have a coefficient correlation of 0.94 with a slope equal to 0.91 in Panel (a). The estimated slope in Panel (a) as well as the comparison of the cumulative distribution functions in Panel (b) suggest that the model performs well in explaining the spatial distribution of workers.

Second, in Panel (c) and (d), I compare the observed local tax revenue to the model-implied local tax revenue by residence. There is a strong positive correlation between the data and the model-implied local tax revenues with a value of 0.92 and an estimated slope of 0.95. In addition, I plot the cumulative distribution functions of the data on local tax revenue and

---

[52]In Appendix C.3, I assess the relationship between the recovered amenities and local outcomes which proxy quality of life at the residential locations (Desmet and Rossi-Hansberg, 2013). Panel (a) of Figure C.4 shows that amenities are higher in places with fewer number of firms discharging wastewater. In Panel (b), residences with lower suicide rates tend to have higher amenities. Lastly, there is a negative correlation between divorce rates and amenities in Panel (c). While I do not formally investigate the relationship between weather and the recovered amenity values as in (Rappaport, 2007), I can infer that nice weather is positively correlated with the recovered amenities because coastal areas tend to have mild weather in Summer and Winter relative to inland districts. Therefore, proximity to the ocean in the coastal districts and its positive relationship with nice weather make coastal districts relatively more attractive.

the model-counterpart. Local government spending is equal to the sum of a fixed fraction of local tax revenues and the intergovernmental transfers, last of which my calibration matches. Therefore, Panel (c) and (d) show that the model explains the spatial distribution of local government spending well.

Third, I verify the model prediction on residential floor space. Panel (e) and (f) compare the residential floor spaces predicted by the model to the observed area of land used for residential purposes measured in $1000m^2$ from the Land Use Statistics in 2015. The correlation coefficient of the two variables is 0.52 and the estimated slope is equal to 0.97. While strong, the relationship between the data and the model-implied values has a relatively low correlation coefficient. This is because the observed data measures total land area used residentially, which does not take the ratio of floor space to land area into account. Despite the sources of measurement error, the model performs well in capturing residential floor spaces.

8   Counterfactual Policy Experiments

In this section, I quantify the welfare consequences of the spatial distribution of local government spending. In particular, I vary the extent of redistribution while holding the rules of redistribution and the extent of fiscal decentralization constant.

8.1   Determinants of Rules of Redistribution

The primary objective of the Local Subsidy Act, which determines the rules of redistribution, is to promote equitable economic growth across localities. As a result, the rules of redistribution is expected to favor residences which are intrinsically less attractive to live (low values of $\tilde{B}_j$) and fiscally weak (low TR_j) to promote economic growth in these districts. It is important to understand the determinants of rules of redistribution because I conduct counterfactual policy experiments while holding the observed rules of redistribution fixed in the subsequent section. I formally study the determinants of the rules of redistribution observed in 2015 in a regression framework. To do so, I regress the log of the observed rules of redistribution $\ln \varsigma_j$ on the log of residential density R_j, recovered amenity values $\tilde{B}_r$, local productivity A_j, and employment density L_j. Table 8 summarizes the estimation results.

In Column (1), the coefficient in front of the log residential density is positive and statistically significant. The result implies that the rules of redistribution is higher in places with higher population density conditional on the geographical area. Introducing the log recovered values of adjusted amenities and productivity in Column (2) and then in Column (3), I find that districts with higher amenity values and productivity receives smaller share of intergovernmental transfers. Lastly, in Column (4), I find that the employment density of a residence does not

affect the rules of redistribution.

8.2 Welfare Consequences of Redistribution

In this section, I conduct a series of counterfactual policy experiments in which I vary the extent of redistribution. Throughout the exercises, I hold the extent of fiscal decentralization (i.e., the fraction of total tax revenue spent locally) constant at the level observed in 2015 $\tilde{\chi} = 0.4$ as well as the rules of redistribution $\{\varsigma_j\}_{j=1}^S$. In each of counterfactuals, I consider varying extent of redistribution denoted by $\tilde{\varsigma}$, which varies from 0 up to $\tilde{\chi}$. Local government spending is expressed as follows:

$$G_j = (\tilde{\chi} - \tilde{\varsigma})TR_j + \varsigma_j\tilde{\varsigma}\sum_{j'=1}^S TR_{j'} \tag{37}$$

If $\tilde{\varsigma} = 0$, local government spending solely depends on local tax revenue. In the other extreme in which $\tilde{\varsigma} = \tilde{\chi}$, intergovernmental transfers completely determine local government expenditures. The observed extent of redistribution is 0.3, which I consider a baseline.

Figure 6 plots the changes in the aggregate welfare of workers $\bar{u}$ as defined in Section 6.1 relative to the baseline level ($\tilde{\varsigma} = 30\%$). When the redistributive intergovernmental transfers are completely eliminated and local government spending is determined solely based on local tax revenue ($\tilde{\varsigma} = 0\%$), the aggregate welfare of workers decrease by 1.2 percent. In the other extreme case in which local government spending is completely determined by intergovernmental transfers ($\tilde{\varsigma} = 40\%$), the aggregate welfare also decreases by 0.3 percent. Considering the varying extent of redistribution (with an increment of 5 percentage points), I find that the aggregate welfare is maximized when the extent of redistribution is equal to 20 percent. This implies that by lowering the extent of redistribution observed in 2015 by 10 percentage points, the aggregate welfare of workers would reach its highest, which is 0.12 percent higher than the baseline level.

The extent of redistribution controls the trade-offs between two types of fiscal spillovers. In districts that are net contributors to redistribution, a dollar tax contribution of a resident is shared with all the other residents living in the same district, but also with other workers living in districts that are net receivers.[53] Therefore, in the presence of redistributive intergovernmental transfers, there are two sources of fiscal spillovers: intra-district and inter-district. The size of intra-district fiscal spillover decreases in the extent of redistribution. It is also necessarily the

---

[53]To help understand the types of spillovers, it is important to reiterate two important characteristics of local government spending in South Korea and more broadly local public finance. First, local government goods and services are not fully rival (i.e., $\theta < 1$). Second, due to redistributive intergovernmental transfers, how much is transferred from districts that are fiscally strong (net contributors) to those with weak fiscal capacities (net receivers) increases in the extent of redistribution. Higher the extent of redistribution, larger the fraction of my tax contribution diverted for redistribution.

case that the size of inter-district fiscal spillover becomes larger as the extent of redistribution increases.

Therefore, the welfare changes summarized in Figure 6 are the consequences of changes in the extents of intra- and inter-district fiscal spillovers. On the one hand, when the extent of redistribution is greater than 20 percent, inter-district spillover serves as a primary source of inefficiency. In this case, intergovernmental transfers raise local government expenditures in net-receiving districts by drawing expenditures from net-contributing districts. In response, workers are attracted to and move to these places which have become less undesirable. On the other hand, when the extent of redistribution is less than 20 percent, intra-district spillover is responsible for lowering the overall welfare. Similarly, in this case, districts that are fiscally strong would attract additional residents from the tax contributions of fellow residents shared within each district.

Figure 7 shows that the extent of fiscal spillovers is minimized when the extent of redistribution is equal to 20 percent. I construct a measure for the extent of fiscal spillovers $\tilde{\varsigma}$ by computing the standard deviation of local government goods and services net of worker tax contribution (i.e., how much extra benefit workers enjoy due to spillovers) for each counterfactual. This measure gauges the dispersion of external benefits of local government spending from intra- and inter-district spillovers. Higher the dispersion, higher the incentives for the workers to reallocate. At the optimum level of redistribution at 20 percent, the extent of fiscal spillovers is reduced by 20 percent.

Lastly, I conduct the same set of counterfactual policy experiments based on two different restrictions commonly imposed in the literature on migration and commuting. First, the migration literature assumes that workers live and work in the same location. I set the semi-elasticity of commuting with respect to commuting distance $\kappa\epsilon$ equal to infinity, the semi-elasticity of migration with respect to migration distance $\rho\epsilon$ equal to 0.007, and the semi-elasticity of job finding to its distance equal to 0. Second, the commuting literature assumes costless migration. Likewise, I assume that the distance-elasticities of migration and job search equal to 0 and the semi-elasticity of commuting to commuting distance equal to 0.074 and compute the counterfactual outcomes. Then, for each of two sets of redistribution separately, I solve for the new equilibrium and compute counterfactual changes in the aggregate worker welfare under vary extents of redistribution.

In Panel (a) of Figure 8, I plot the welfare changes relative to the baseline in 2015 assuming no inter-district commuting. If workers are not allowed to commute outside of districts, eliminating redistribution altogether leads in a higher welfare loss of about 2 percent. Furthermore, the optimal extent of redistribution is higher at a level close to 30 percent. Workers are not able to access districts with higher productivity without moving into these districts. Then, workers agglomerate in these districts, contributing to increasing intra-district fiscal spillover.

As a result, there is a demand for greater redistribution.

Panel (b) of Figure 8 plots the changes in the worker welfare under the assumption of no migration and job finding costs. While not optimal, eliminating redistributive intergovernmental transfers lead to a sizable increase in welfare by about 2.3 percent. This implies that the need for redistribution across districts is small when workers can migrate across districts freely. With no migration and job finding costs, it becomes easier for workers to access districts with higher productivity. At the same time, in the presence of redistributive intergovernmental transfers, workers find it profitable to reside in net-receiving districts with positive inter-district fiscal spillovers at the expense of longer commute because they benefit from local government goods and services more than their tax contribution. Therefore, with no migration and job finding costs, lowering the extent of redistribution increases the overall efficiency of the economy.

9   Conclusion

In this paper, I make three contributions to our understanding of local provision of government goods and services as a determinant of the spatial distribution of workers. First, I present a quantitative general equilibrium in which workers make both migration and commuting decisions, which have been traditionally studied separately. The key prediction of the model is a gravity equation summarizing the distribution of workers in terms of three locations: previous residence, current residence, and workplace.

Second, I combine the framework with the quasi-natural experiment leading to plausibly exogenous variation in local spending and estimate the key reduced-form elasticities of worker mobility with respect to local government expenditure, residential density, and home prices. In addition, I estimate the elasticities of worker mobility with respect to spatial frictions (migration, commuting, and job finding). The key finding is that the marginal valuation of local government spending is equal to 75 cents of after-tax income. I show that there are large biases when the elasticity of migration with respect to distance is estimated without accounting for commuting patters and vice versa. The results altogether show that where workers lived before matters for not only where they live today, but only where they presently work.

Third, based on counterfactual policy experiments, I show that there exists a fiscal arrangement (local taxation vs. intergovernmental transfers), which would maximize the overall welfare of workers. I discuss how the optimal mix of local taxation and intergovernmental transfers balances the trade-offs between the extents of intra-regional and inter-regional fiscal spillovers. The results suggest that reducing the current extent of redistribution, thereby allowing local governments to rely more on their local income tax, would increase the overall efficiency. Furthermore, spatial frictions and their effects on worker mobility are important in determining an

optimal fiscal arrangement. If workers are assumed to live and work in the same location (the key assumption in the migration literature), the importance of redistributive intergovernmental transfers are overemphasized. However, if workers can migrate without any costs (the key assumption in the commuting literature), the importance of local taxation (less redistribution) is overemphasized.

Overall, I find that it is crucial to account for both margins of mobility (i.e., migration and commuting) not only to understand the determinants and their effects on the spatial distribution of workers and more broadly economic activity, but also to inform policy makers of the welfare consequences from the spatial distribution of local government spending.

References

Adao, Rodrigo, Michal Kolesár, and Eduardo Morales, "Shift-Share Designs: Theory and Inference," *The Quarterly Journal of Economics*, Forthcoming, pp. 1–44.

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf, "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*, 2015, *83* (6), 2127–2189.

Ákao Valentinyi and Berthold Herrendorf, "Measuring Factor Income Shares at the Sectoral Level," *Review of Economic Dynamics*, 2008, *11*, 820–835.

Akcigit, Ufuk, Salome Baslandze, and Stefanie Stantcheva, "Taxation and the International Mobility of Inventors," *American Economic Review*, 2016, *106* (10), 2930–2981.

Albert, Christoph and Joan Monras, "Immigration and Spatial Equilibrium: the Role of Expenditures in the Country of Origin," *Working Paper*, 2019, pp. 1–86.

Albouy, David, "The Unequal Geographic Burden of Federal Taxation," *Journal of Political Economy*, 2009, *117* (4), 635–667.

_ , "Evaluating the Efficiency and Equity of Federal Fiscal Equalization," *Journal of Public Economics*, 2012, *96*, 824–839.

Balboni, Clare, "In Harm's Way? Infrastructure Investments and the Persistence of Coastal Cities," *Working Paper*, 2019, pp. 1–86.

Bartik, Timothy J., *Who Benefits from State and Local Economic Development Policies?*, Upjohn Press, W.E. Upjohn Institute for Employment Research, 1991.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How Much Should We Trust Differences-In-Differences Estimates?," *The Quarterly Journal of Economics*, February 2004, *119* (1), 249–275.

Bianchi, Nicola, Michela Giorcelli, and Enrica Maria Martino, "The Effects of Fiscal Decentralization on Publicly Provided Services and Labor Markets," *Working Paper*, 2019, pp. 1–40.

Black, Sandra E., "Do Better Schools Matter? Parental Valuation of Elementary Education," *The Quarterly Journal of Economics*, May 1999, *114* (2), 577–599.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, "Quasi-Experimental Shift-Share Research Designs," *NBER Working Paper No. 24997*, 2018, pp. 1–41.

Bryan, Gharad and Melanie Morten, "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia," *Journal of Political Economy*, 2018, pp. 1–78.

Cadena, Brian and Brian Kovak, "Immigrants Equilibrate Local Labor Markets: Evidence from the Great Recession," *American Economic Journal: Applied Economics*, 2016, *8* (1), 257–290.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro, "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, May 2019, *87* (3), 741–835.

Card, David, "Immigration Inflows, Native Outflows, and the Local Market Impacts of Higher Immigration," *Journal of Labor Economics*, 2001, *19* (1), 22–64.

Cellini, Stephanie R., Fernando Ferreira, and Jesse Rothstein, "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design," *The Quarterly Journal of Economics*, 2010, *125* (1), 215–261.

Chay, Kenneth Y. and Michael Greenstone, "Does Air Quality Matter? Evidence from the Housing Market," *Journal of Political Economy*, April 2005, *113* (2), 376–424.

Ciccone, Antonio and Robert Hall, "Productivity and the Density of Economic Activity," *American Economic Review*, 1996, *86* (1), 54–70.

Combes, Pierre and Laurent Gobillon, "The Empirics of Agglomeraion Economics," in Gilles Duranton, Vernon Henderson, and William Strange, eds., *Handbook of Regional and Urban Economics*, Amsterdam: Elsevier, 2015, chapter 5, pp. 247–348.

Congress of South Korea, "Local Tax Act (No.14033)," Online September 2016.

Davis, Marris A. and François Ortalo-Magné, "Household Expenditures, Wages, Rents," *Review of Economic Dynamics*, 2011, *14*, 248–261.

de la Roca, Jorge and Diego Puga, "Learning by Working in Big Cities," *Review of Economic Studies*, 2017, *84*, 106–142.

Desmet, Klaus and Esteban Rossi-Hansberg, "Urban Accounting and Welfare," *American Economic Review*, 2013, *103* (6), 2296–2327.

Diamond, Rebecca, "The determinants and welfare implications of US workers' diverging location choices by skill: 1980 to 2000," *American Economic Review*, 2016, *106* (3), 479–524.

Dix-Carneiro, Rafael, "Trade Liberalization and Labor Market Dynamics," *Econometrica*, May 2014, *82* (3), 825–885.

Donaldson, Dave and Richard Hornbeck, "Railroads and American Economic Growth: A Market Access Approach," *Quarterly Journal of Economics*, 2016, *131* (2), 799–858.

Ducruet, César, Réka Juhász, Dávid Krisztián Nagy, and Claudia Steinwender, "All Aboard: The Aggregate Effects of Port Development," *Working Paper*, 2019, pp. 1–70.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, 2002, *70* (5), 1741–1779.

Eeckhout, Jan, Roberto Pinheiro, and Kurt Schmidheiny, "Spatial Sorting," *Journal of Political Economy*, 2014, *122* (3), 554–620.

Fajgelbaum, Pablo D. and Cecile Gaubert, "Optimal Spatial Policies, Geography and Sorting," *NBER Working Paper 24632*, 2018, pp. 1–64.

_ , **Eduardo Morales, Juan Carlos Suarez-Serrato, and Owen Zidar**, "State Taxes and Spatial Misallocation," *Review of Economic Studies*, 2019, *86* (1), 333–376.

Fisman, Raymond and Roberta Gatti, "Decentralization and Corruption: Evidence from Countries," *Journal of Public Economics*, 2002, *83*, 325–345.

Gelbach, Jonah B., "Migration, the Life Cycle, and State Benefits: How Low Is the Bottom?," *Journal of Political Economy*, 2004, *112* (5), 1091–1130.

Glaeser, Edward and David Maré, "Cities and Skills," *Journal of Labor Economics*, 2001, *19* (2), 316–342.

_ **and Joshua Gottlieb**, "The Economics of Place-making Policies," *Brookings Papers on Economic Activity*, 2008, *Spring*, 155–239.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik Instruments: What, When, Why, and How," *NBER Working Paper 24408*, 2018, pp. 1–66.

Kleven, Henrik Jacobsen, Camille Landais, Emmanuel Saez, and Esben Schultz, "Migration and Wage Effects of Taxing Top Earners: Evidence from the Foreigners' Tax Scheme in Denmark," *The Quarterly Journal of Economics*, 2014, pp. 333–378.

Kline, Patrick and Enrico Moretti, "People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs," *The Annual Review of Economics*, 2014, *6*, 629–662.

Lee, Jungmin, "The Labor Market in South Korea, 2000-2016," Technical Report, IZA 2017.

Manning, Alan and Barbara Petrongolo, "How Local Are Labor Markets? Evidence from a Spatial Job Search Model," *American Economic Review*, 2017, *107* (10), 2877–2907.

Marinescu, Ioana and Roland Rathelot, "Mismatch Unemployment and the Geography of Job Search," *American Economic Journal: Macroeconomics*, July 2018, *10* (3), 42–70.

McFadden, Daniel, "The Measurement of Urban Travel Demand," *Journal of Public Economics*, 1974, *3* (4), 303–328.

Molloy, Raven, Christopher L. Smith, and Abigail Wozniak, "Internal Migration in the United States," *Journal of Economic Perspective*, 2011, *25* (3), 173–196.

Monte, Ferdinando, Stephen Redding, and Esteban Rossi-Hansberg, "Commuting, Migration, and Local Employment Elasticities," *American Economic Review*, 2018, *108* (12), 2855–3890.

Moretti, Enrico and Daniel J. Wilson, "The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists," *American Economic Review*, 2017, *107* (7), 1858–1903.

Morten, Melanie and Jacqueline Oliveira, "The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City," *NBER Working Paper 22158*, 2018, pp. 1–64.

Moulton, Brent R., "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units," *The Review of Economics and Statistics*, May 1990, *72* (2), 334–338.

Oates, Wallace E., "An Essay on Fiscal Federalism," *Journal of Economic Literature*, 1999, *37* (3), 1120–1149.

OECD, "Better Life Index: Korea," Technical Report 2016.

_ , "General Government Spending (Indicator)," Technical Report, OCED 2019.

Pellegrina, Heitor S. and Sebastian Sotelo, "Migration, Specialization, and Trade: Evidence from the Brazilian March to the West," *Working Paper*, June 2019, pp. 1–43.

Piyapromdee, Suphanit, "The Impact of Immigration on Wages, Internal Migration, and Welfare," *Working Paper*, 2017, pp. 1–83.

Preston, Samuel, Patrick Heuveline, and Michel Guillot, *Demography: Measuring and Modeling Population Processes*, Wiley-Blackwell, October 2000.

Rappaport, Jordan, "Moving to Nice Weather," *Regional Science and Urban Economics*, 2007, *37*, 375–398.

Redding, Stephen and Esteban Rossi-Hansberg, "Quantitative Spatial Economics," *Annual Review of Economics*, 2017, *9*, 21–58.

_ **and Matthew Turner**, "Transportation Costs and the Spatial Orgnization of Economic Activity," in Gilles Duranton, Vernon Henderson, and William Strange, eds., *Handbook of Regional and Urban Economics*, Amsterdam: Elsevier, 2015, chapter 20, pp. 1339–1398.

Roback, Jennifer, "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, 1982, *90* (6), 1257–1278.

Rosen, Sherwin, "Wages-based Indexes of Urban Quality of Life," *Current Issues in Urban Economics*, 1979, pp. 34–55.

Sanderson, Eleanor and Frank Windmeijer, "A Weak Instrument F-Test in Linear IV Models with Multiple Endogenous Variables," *Journal of Econometrics*, 2016, *190*, 212–221.

Schafer, Andreas, "Regularities in Travel Demand: an International Perspective," *Journal of Transportation Statistics*, 2000, *3*, 1–31.

Stock, James, Jonathan Wright, and Motohiro Yogo, "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business & Economic Statistics*, 2002, *20* (4), 518–529.

Suárez-Serrato, Juan Carlos and Philippe Wingender, "Estimating the Incidence of Government Spending," *Working Paper*, 2014, pp. 1–69.

_ **and** _ , "Estimating Local Fiscal Multipliers," *Working Paper*, 2016, pp. 1–99.

Tiebout, Charles, "A Pure Theory of Local Expenditures," *Journal of Political Economy*, 1956, *64* (5), 416–424.

Tsivanidis, Nick, "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogotá's TransMilenio," *Working Paper*, 2019, pp. 1–56.

Wildasin, David, "Locational Efficiency in a Federal System," *Regional Science and Urban Economics*, November 1980, *10*, 453–471.

Wildasin, David E., "Theoretical Analysis of Local Public Economics," *Handbook of Regional and Urban Economics*, 1987, *2*, 1131–1178.

Tables

Table 1: Summary Statistics

Variable	(1) Observation	(2) Mean	(3) Std. Dev.	(4) Min	(5) Max
A. Commuting Patterns					
Commuters from Residence	666	0.274	0.236	0	0.773
Commuters to Workplace	666	0.298	0.210	0	0.916
B. Migration Patterns					
Migrants to Residence	666	0.187	0.079	0.053	0.559
Out-Migrants from Residence	666	0.180	0.071	0.048	0.443
C. Local Government Budget					
Total Local Expenditure	666	362,785	233,524	59,614	1,881,082
Per-Capita Local Expenditure	666	7.638	5.752	0.904	29.622
Local Income Tax Revenue	666	64,067	95,793	4,020	779,143
Intergovernmental Transfers	666	242,517	128,348	109,239	799,009

Notes: In this table, I report summary statistics computed based on 222 districts in 2005, 2010, and 2015. The data used for Panel A and B is constructed from the Population Census of South Korea. Variable *Commuters from Residence* measures the fraction of residents commuting outside of their district of residency. Variable *Commuters to Workplace* measures the fraction of workers employed in a district who commute from other districts. Similarly, Variable *Migrants to Residence* and *Out-Migrants from Residence* measure the fraction of residents in a district who moved in within 5 years and the fraction of residents who moved out of a district within 5 years. Panel C is computed using the Yearbook of Local Public Finance data. The unit for the values reported in Panel (c) is 1 million KRW (approximately 1,000 USD). See Section 2.1 for the details on the data sources.

Table 2: (OLS) Elasticities of Worker Mobility with respect to Local Government Goods

Dependent Variable:	(1) $\ln \pi_{orm,t}$	(2) $\ln \pi_{orm,t}$	(3) $\ln \pi_{orm,t}$	(4) $\ln \pi_{orm,t}$
Local Government Expenditure, $\ln G_{r,t}$ ($\beta_G = \lambda\epsilon$)	-0.231***	0.0965**	0.101	0.0957***
	(0.0149)	(0.0405)	(0.141)	(0.0299)
Number of Households, $\ln R_{r,t}$ ($\beta_R = \theta\lambda\epsilon$)	0.120***	0.0608**	0.297	0.590***
	(0.0125)	(0.0268)	(0.218)	(0.0522)
Floor Space Price, $\ln Q_{r,t}$ ($\beta_Q = (1-\beta)\epsilon$)	-0.0416***	-0.0101	0.00802	-0.00148
	(0.0129)	(0.0334)	(0.0336)	(0.00653)
Observations	258,323	258,323	258,323	258,323
Fixed Effects:				
Job Finding Pair $\times$Year ($\phi_{om,t}$)	N	Y	Y	Y
Migration Pair (ϕ_{or})	N	N	Y	Y
Commuting Pair (ϕ_{rm})	N	N	N	Y

Notes: In this table, I report the OLS estimates of elasticities of worker's mobility to local government expenditure and resident population levels based on Equation 6, starting with a simple estimate without any fixed effects in Column (1) and gradually adding the fixed effects discussed in Section 4.1.1. Column (4) corresponds to Equation 6 with the full set of fixed effects. The sample is from 3 waves of the Population Census of South Korea in 2005, 2010, and 2015, based on 3,500,232 male household heads who are employed between the ages of 25 and 60. Each observation corresponds to a triplet of previous and current residences and workplace location. Robust standard errors in parentheses, with multi-way clustering by migration pair $\times$ year, commuting pair $\times$ year, and a triplet of previous and current residences and workplace: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table 3: (2SLS) Elasticities of Worker Mobility with respect to Local Government Goods

Dependent Variable:	(1) OLS $\ln \pi_{orm,t}$	(2) First Stage $\ln G_{r,t}$	(3) First Stage $\ln R_{r,t}$	(4) First Stage $\ln Q_{r,t}$	(5) 2SLS $\ln \pi_{orm,t}$
Local Government Expenditure, $\ln G_{r,t}$ ($\beta_G = \lambda\epsilon$)	0.0957***				1.072***
	(0.0299)				(0.387)
Number of Households, $\ln R_{r,t}$ ($-\beta_R = -\theta\lambda\epsilon$)	0.590***				-0.844
	(0.0522)				(0.622)
Floor Space Prices, $\ln Q_{r,t}$ ($-\beta_Q = -(1-\beta)\epsilon$)	-0.00148				-0.490***
	(0.00653)				(0.067)
Predicted Tax Contribution (low), $IV_{r,t}^{low}$		13.26***	5.592***	60.23***	
		(1.518)	(0.246)	(6.851)	
Predicted Tax Contribution (high), $IV_{r,t}^{high}$		13.921***	6.761***	21.30***	
		(0.833)	(0.246)	(3.257)	
Number of Households 30 years ago, $IV_{r,t}^{R}$		0.028***	-0.014**	0.110***	
		(0.009)	(0.007)	(0.030)	
Observations	258,323	258,323	258,323	258,323	258,323
SW F-stat		19.01	16.87	25.66	
$\hat{\theta}$	-6.164**				0.787**
	(2.176)				(0.315)

Notes: In this table, I compare the OLS estimates and 2SLS estimates of elasticities of worker's mobility to local government expenditure and resident population levels based on Equation 6. Column (1) is identical to Column (4) in Table 2. Column (2) and Column (3) report the first stage results. The 2SLS estimates are reported in Column (4). Across columns, the full set of fixed effects as discussed in Section 4.1.1 are included. The sample ($N = 258,323$) is from 3 waves of the Population Census of South Korea in 2005, 2010, and 2015, based on 3,500,232 male household heads who are employed between the ages of 25 and 60. Each observation corresponds to a triplet of previous and current residences and workplace location. Robust standard errors for Column (1), (2), and (3) and bootstrapped (20,000 replications) standard errors for Column (4) in parentheses, with multi-way clustering by migration pair × year, commuting pair × year, and a triplet of previous and current residences and workplace: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$. I estimate the congestion parameter θ based on the structural relationship between the estimated reduced form parameters ($\beta_R/\beta_G = \theta$): $\hat{\theta} = 0.787$ (0.199).

52

Table 4: Semi-Elasticity of Migration with respect to Distance

Dependent Variable:	(1) $\ln \pi_{orm,t}$	(2) $\ln \pi_{orm,t}$	(3) $\ln \pi_{orm,t}$	(4) $\ln \pi_{orm,t}$	(5) $\ln \pi_{or,t}$
distance d_{or} $(-\rho\epsilon)$	-0.002***	-0.004***	-0.033***	-0.033***	-0.007***
	(0.0001)	(0.0001)	(0.0009)	(0.001)	(0.0001)
distance×2005				0.0001	
				(0.002)	
distance×2010				0.0005	
				(0.002)	
Fixed effects:					
Commute Pair ×Year $(\phi_{rm,t})$	N	Y	Y	Y	N
Job Finding Pair ×Year $(\phi_{om,t})$	N	N	Y	Y	N
Origin ×Year $(\phi_{o,t})$	N	N	N	N	Y
Current Residence × Year $(\phi_{r,t})$	N	N	N	N	Y

Notes: In this table, I estimate the semi-elasticity of migration with respect to distance based on Equation 14, starting with a simple estimate without any fixed effects in Column (1) and gradually adding the fixed effects. Column (3) corresponds to Equation 14. Column (4) tests whether the semi-elasticity is time-invariant or not. In Column (5), I report the estimated coefficient based on Equation (15) following the literature on migration. The sample is from 3 waves of the Population Census of South Korea in 2005, 2010, and 2015, based on 3,500,232 male household heads who are employed between the ages of 25 and 60. Each observation corresponds to a triplet of previous and current residences and workplace location for Columns (1) - (4). Robust standard errors in parentheses, with multi-way clustering by migration pair × year, commuting pair × year, job finding pair × year, and a triplet of previous and current residences and workplace for Columns (1) - (4) : $***p < 0.01$, $**p < 0.05$, $* < 0.1$. The unit of observation for Column (5) is a pair of previous and current residences. Robust standard errors in parentheses, with three-way clustering by previous residence × year, current residence × year, and migration pair for Column (5): $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table 5: Semi-Elasticity of Commuting with respect to Distance

Dependent Variable:	(1) $\ln \pi_{orm,t}$	(2) $\ln \pi_{orm,t}$	(3) $\ln \pi_{orm,t}$	(4) $\ln \pi_{orm,t}$	(5) $\ln \pi_{rm,t}$
distance d_{rm} $(-\kappa\epsilon)$	-0.013***	-0.035***	-0.045***	-0.046***	-0.074***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
distance$\times$2005				0.001	
				(0.001)	
distance$\times$2010				0.003**	
				(0.001)	
Observations	258,323	258,323	258,323	258,323	20,676
Fixed effects:					
Migration Pair $\times$Year $(\phi_{or,t})$	N	Y	Y	Y	N
Job Finding Pair $\times$Year $(\phi_{om,t})$	N	N	Y	Y	N
Current Residence $\times$Year $(\phi_{r,t})$	N	N	N	N	Y
Workplace $\times$Year $(\phi_{m,t})$	N	N	N	N	Y

Notes: In this table, I estimate the semi-elasticity of commuting with respect to distance based on Equation (16), starting with a simple estimate without any fixed effects in Column (1) and gradually adding the fixed effects. Column (3) corresponds to Equation (16). Column (4) tests whether the semi-elasticity is time-invariant or not. In Column (5), I report the estimated coefficient based on Equation (17) following the literature on commuting. The sample is from 3 waves of the Population Census of South Korea in 2005, 2010, and 2015, based on 3,500,232 male household heads who are employed between the ages of 25 and 60. Each observation corresponds to a triplet of previous and current residences and workplace location for Columns (1) - (4). Robust standard errors in parentheses, with multi-way clustering by migration pair $\times$ year, commuting pair $\times$ year, job finding pair $\times$ year, and a triplet of previous and current residences and workplace for Columns (1) - (4): $***p < 0.01, **p < 0.05, * < 0.1$. The unit of observation for Column (5) is a pair of current residence and workplace location. Robust standard errors in parentheses, with three-way clustering by current residence $\times$ year, workplace location $\times$ year, and commuting pair for Column (5): $***p < 0.01, **p < 0.05, *p < 0.1$.

Table 6: Semi-Elasticity of Job Finding with respect to Distance

Dependent Variable:	(1) $\ln \pi_{orm,t}$	(2) $\ln \pi_{orm,t}$	(3) $\ln \pi_{orm,t}$	(4) $\ln \pi_{orm,t}$
distance d_{om} (-$\delta\epsilon$)	-0.001***	-0.004***	-0.016***	-0.015***
	(0.00001)	(0.00001)	(0.0003)	(0.0004)
distance×2005				-0.002**
				(0.001)
distance×2010				0.001
				(0.001)
Observations	258,323	258,323	258,323	258,323
Fixed effects:				
Commute Pair ×Year ($\phi_{rm,t}$)	N	Y	Y	Y
Migration Pair ×Year ($\phi_{or,t}$)	N	N	Y	Y

Notes: In this table, I estimate the semi-elasticity of commuting with respect to distance based on Equation (18), starting with a simple estimate without any fixed effects in Column (1) and gradually adding the fixed effects. Column (3) corresponds to Equation (18). Column (4) tests whether the semi-elasticity is time-invariant or not. The sample is from 3 waves of the Population Census of South Korea in 2005, 2010, and 2015, based on 3,500,232 male household heads who are employed between the ages of 25 and 60. Each observation corresponds to a triplet of previous and current residences and workplace location. Robust standard errors in parentheses, with multi-way clustering by migration pair × year, commuting pair × year, job finding pair × year, and a triplet of previous and current residences and workplace: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table 7: Summary of Parameterization

Parameter	Description	Value	Method	Source
α	labor share	0.823	estimated	Economic Census
$1-\beta$	housing expenditure share	0.15	estimated	HH Expenditure Survey
ϵ	shape parameter	3.54	estimated	fixed effects
λ	value of local gov't goods	0.303	estimated	Gravity Equation
θ	net congestion	0.787	estimated	Gravity Equation
ρ	spatial decay of migration	0.009	estimated	Gravity Equation
κ	spatial decay of commuting	0.013	estimated	Gravity Equation
δ	spatial decay of job finding	0.005	estimated	Gravity Equation
τ	income tax rate	0.245	observed	Income Tax Act
ς	local-national revenue sharing	0.091	observed	Local Tax Act
χ	extent of redistribution	0.35	observed	Local Subsidy Act
$\{\varsigma_j\}$	redistribution		observed	Local Subsidy Act
$\{A_j\}$	productivity		recovered	PM+ZP
$\{\tilde{B}_j\}$	adjusted amenities		recovered	fixed effects
$\{H_j\}$	floor space		recovered	Floor space market clearing

Notes: This table summarizes the estimates of the structural parameters of the model. Note that I estimate the value of $1-\beta$ using the Household Expenditure Survey of 2015. Alternatively, based on the estimation result summarized in Table 3 and the estimated value of ϵ, I can recover the structural value of $1-\beta$ equal to 0.14. See Section 7 for a detailed description of how the parameters above are estimated and recovered.

Table 8: Determinants of Redistribution Policy in 2015

Dependent Variable	(1)	(2)	(3)	(4)
	Observed Redistribution Policy ($\ln \varsigma_r$)			
Residential Population ($\ln R_r$)	0.441***	0.530***	0.610***	0.631***
	(0.0269)	(0.0285)	(0.0284)	(0.0578)
Amenities ($\ln \tilde{B}_r$)		-0.170***	-0.184***	-0.187***
		(0.0323)	(0.0255)	(0.0273)
Productivity ($\ln A_r$)			-0.636***	-0.605***
			(0.0906)	(0.130)
Employment Population ($\ln L_r$)				-0.0296
				(0.0730)
Area ($\ln Area_r$)	0.259***	0.262***	0.253***	0.254***
	(0.0146)	(0.0126)	(0.0113)	(0.0111)
Observations	222	222	222	222
R^2	0.619	0.668	0.744	0.744

Notes: In this table, I investigate the determinants of the rules of redistribution by regressing the log of percentage of the total intergovernmental transfers each district receives against local characteristics. The dependent variation is the log of the share of intergovernmental transers each district recieved in 2015. I begin with covariates of residential population and area Column (1) and gradually introduce additional covariates across columns. Each observation correspond to a district in 2015.

Figures

Figure 1: Commuting and Migration Patterns vs. Distance

(a) Commuting
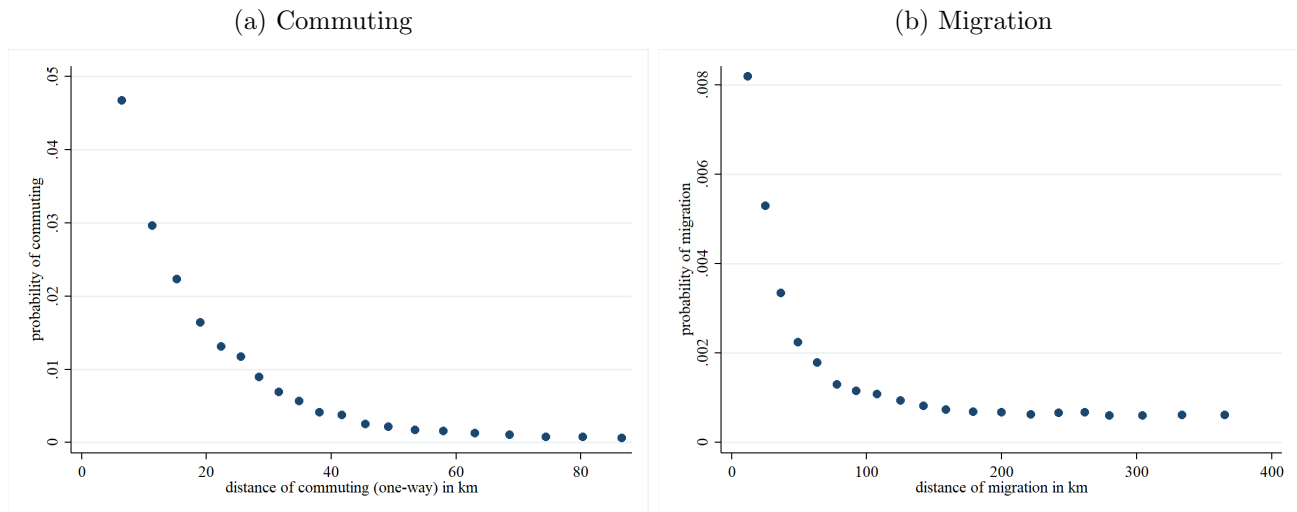
(b) Migration

Notes: This figure shows that the probabilities of commuting shown in Panel (a) and the probabilities of migration shown in Panel (b) decrease as distances of commuting and migration increase. The probabilities of commuting and migration are computed using the Population Census of South Korea (2005, 2010, and 2015). Each point corresponds to 5 percentiles of commuting and migration distances.

Figure 2: Spatial Distribution of Residential Density and Local Government Spending

(a) Residential Density

(b) Local Government Spending
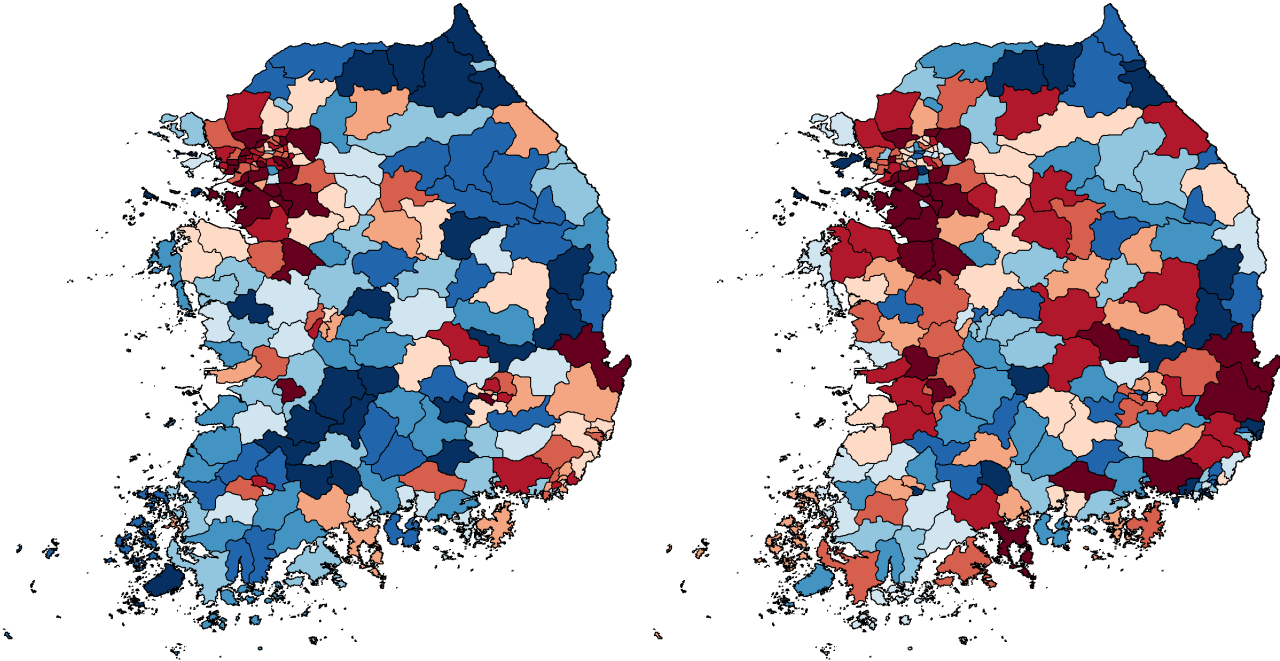
Notes: The figure on the left plots the spatial distribution of workers in terms of their residences in 2015. The figure on the right plots the spatial distribution of local government spending in 2015. Red (blue) districts indicate higher (lower) values.

Figure 3: Workers appear willing to migrate/commute longer with higher government spending and lower housing prices

(a) Migration Distance vs. Local Gov't Spending
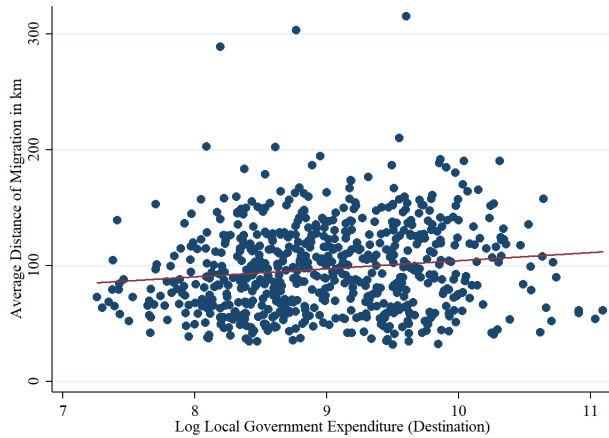
(b) Commuting Distance vs. Local Gov't Spending
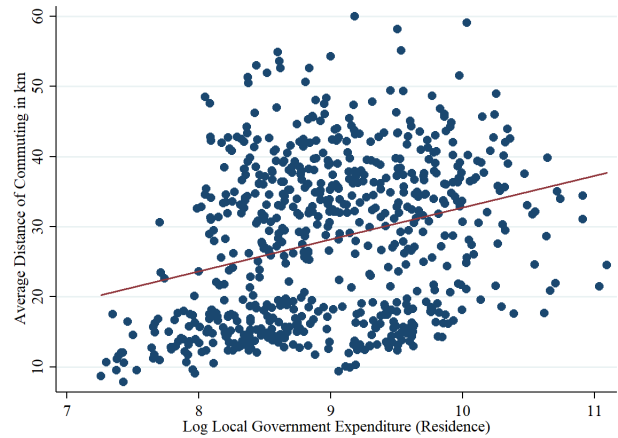
(c) Migration Distance vs. Home Prices
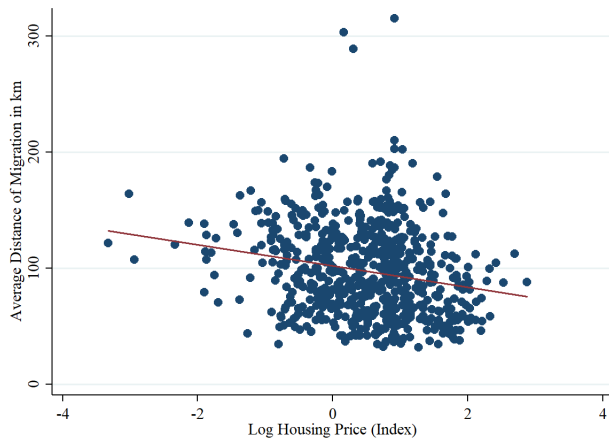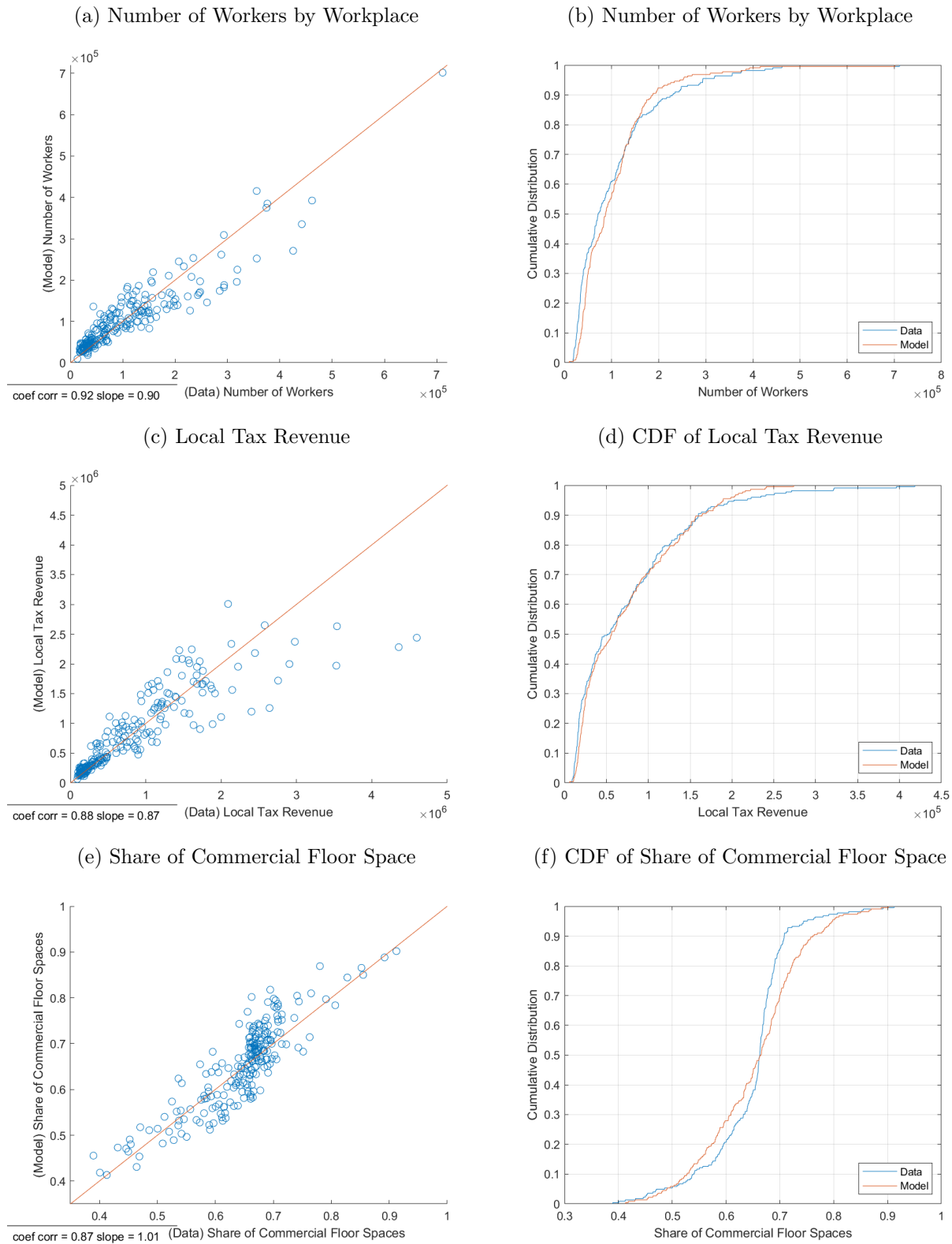
(d) Commuting Distance vs. Housing Prices
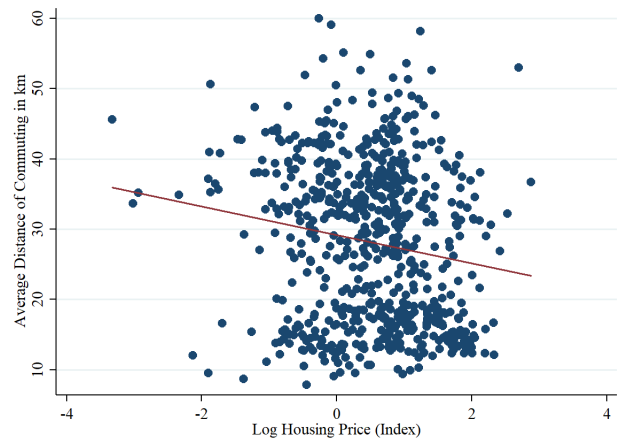
Notes: This figure shows the raw correlation between how far workers migrate and commute and local government spending/Housing prices. Each observation is a district-year pair. The figures in the left plot the average distance that residents have migrated over the past 5 years against local government spending in Panel (a) and against home prices in Panel (c). The figure in the right plot the average distance of commuting for a resident for each district against local government spending in Panel (b) and against the home prices in Panel (d).

Figure 4: Marginal Income Tax Rates before and after 2008 and 2012
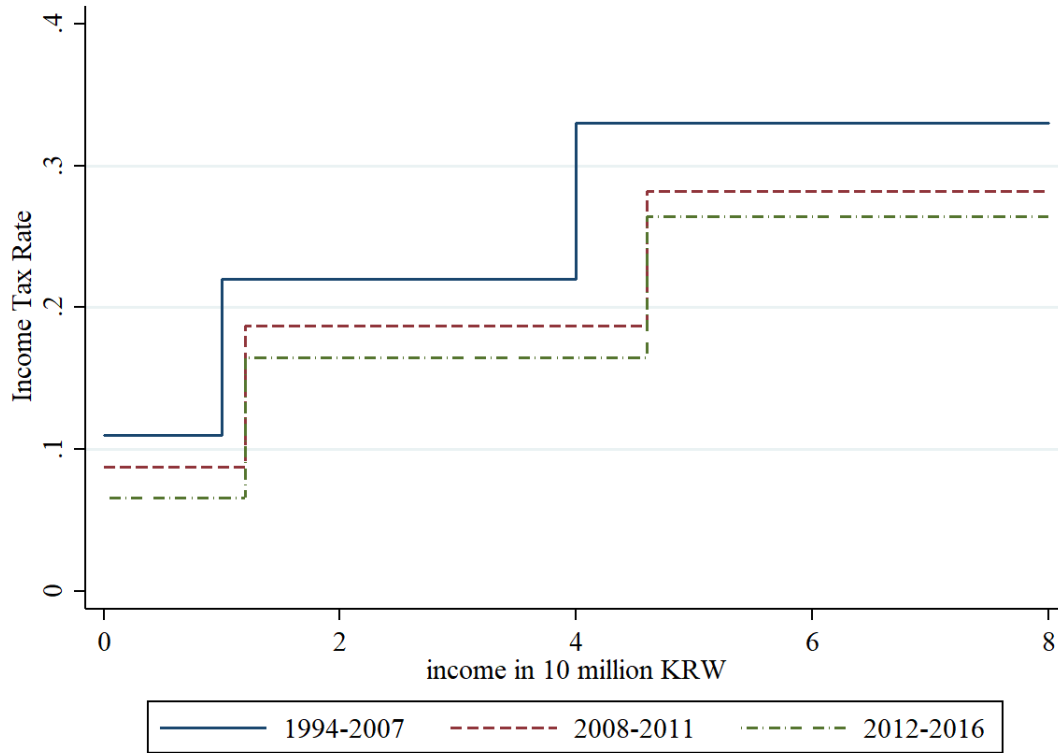
Notes: This figure plots the progressive income tax rates against income measured in 10 million KRW (approximately 10,000 USD) before and after the two episodes of national tax policy reforms in 2008 and 2012. The national income tax rates are outlined in the Income Tax Act. Note the median after-tax income in South Korea in 2015 is 18,180 USD.

Figure 5: Over-identifying Moments: Model vs. Data

(a) Number of Workers by Workplace

(b) Number of Workers by Workplace

(c) Local Tax Revenue

(d) CDF of Local Tax Revenue

(e) Share of Commercial Floor Space

(f) CDF of Share of Commercial Floor Space

Notes: This figure compares 2015 data with model predictions of non-targeted moments. Panel (a) and (b) plot the spatial distribution of workers by employment location. Panel (c) and (d) plot local tax revenues collected at each residence measured in 1 million KRW. Panel (e) and (f) plot the shares of commercial floor space. The straight lines in Panel (a), (c), and (e) are 45 degree lines.

62

Figure 6: Aggregate Welfare Changes and Redistribution
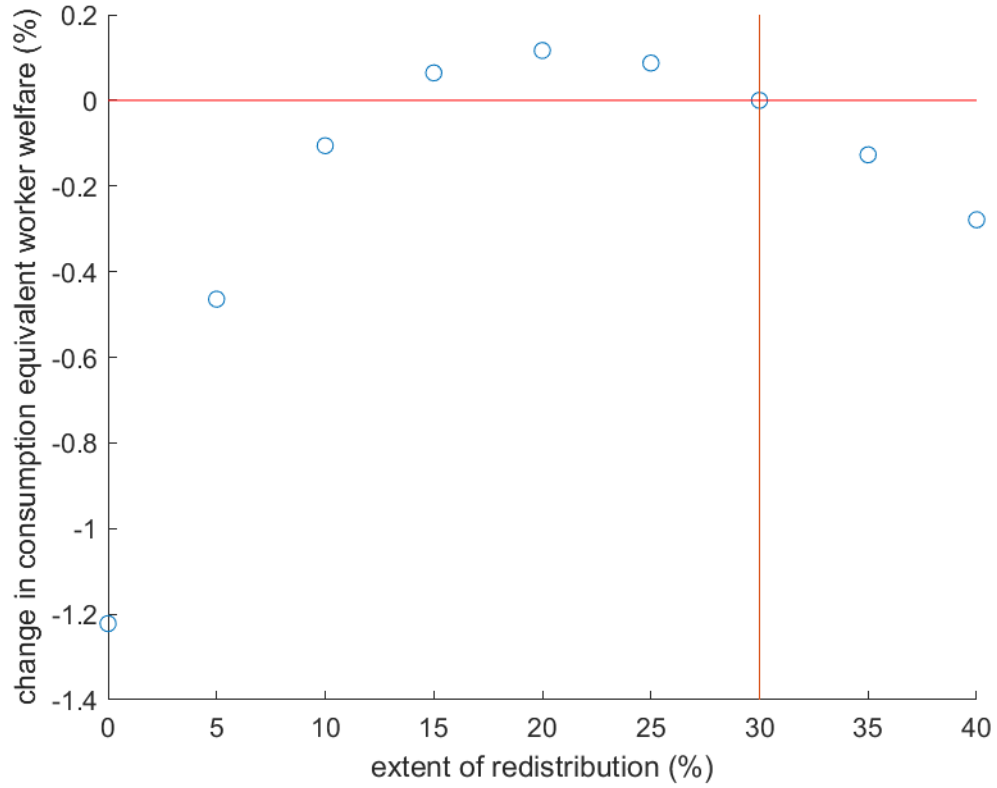
Notes: This figure plots the changes in the aggregate consumption equivalent worker welfare relative to the welfare level in baseline in which the extent of redistribution is equal to 0.3, the observed level in 2015.

Figure 7: Changes Extent of Fiscal Spillover Changes and Redistribution
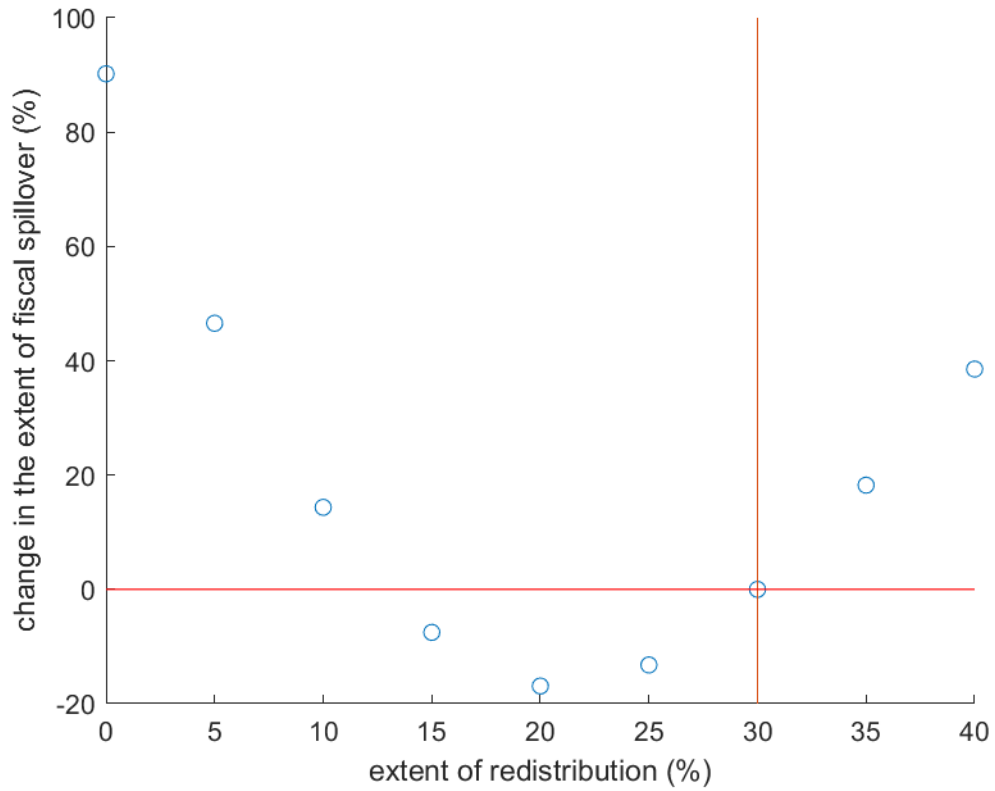
Notes: This figure plots the changes in extent of fiscal spillover to the baseline level in which the extent of redistribution is equal to 0.3, the observed level in 2015. The extent of fiscal spillover measures the dispersion of local government goods and services net of individual tax contribution (i.e., how much extra benefit workers enjoy due to spillovers) across districts.

Figure 8: Aggregate Welfare Changes under Alternative Assumptions

(a) Prohibitively Costly Commuting   (b) Costless Migration and Job Finding
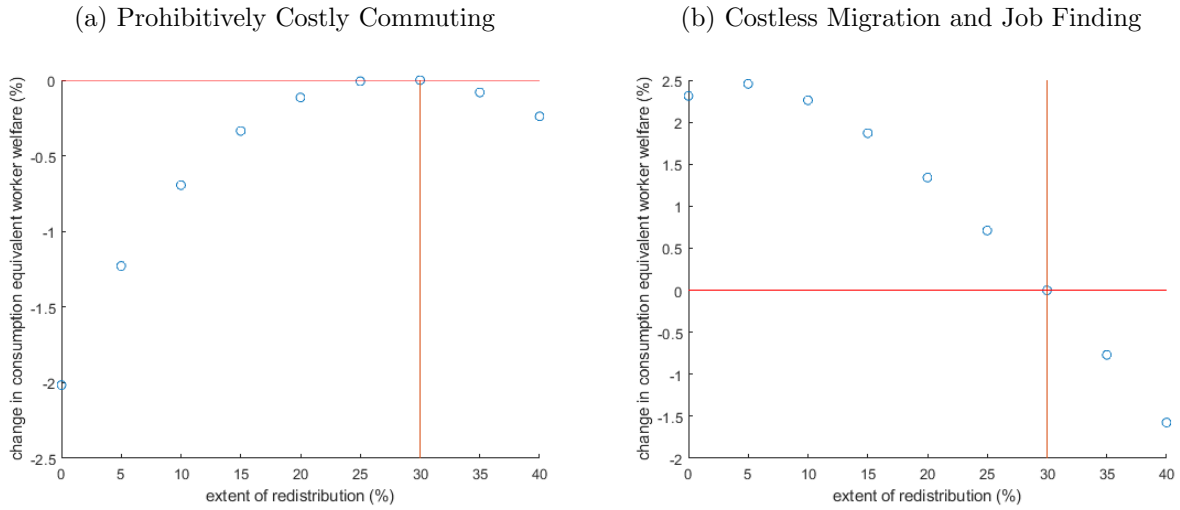
Notes: In this figure, I plot the changes in the consumption equivalent welfare of workers based on two alternative assumptions about spatial frictions. First, I follow the common spatial redistribution imposed in the migration literature (i.e., workers cannot work outside of their district of residence). I solve for a new equilibrium for 2015 assuming the distance elasticity of migration equal to 0.007 as in Column (5) of Table 4, of commuting equal to ∞, and of job search equal to 0. I compute the counterfactual outcomes and plot the changes in worker welfare relative to 2015 (extent of redistribution = 30%) in Panel (a). Second, I follow the common spatial redistribution imposed in the commuting literature (i.e., there is no bilateral cost of commuting and job search). I solve for a new equilibrium for 2015 assuming the distance elasticity of migration equal to 0, of commuting equal to 0.074 as in Column (5) of Table 5, and of job search equal to 0. I compute the counterfactual outcomes and plot the changes in worker welfare relative to 2015 (extent of redistribution = 30 percent) in Panel (b).

65

Appendix

A   Data Appendix

Wages

I construct wages for each district based on the Economic Census of South Korea in 2015. The Census surveys the universe of establishments in South Korea and records the number of employees and the total costs of labor. I aggregate these two information across establishments in each district and divide the total costs of labor by the number of employees to obtain the district-level wages.

Floor Space Prices

The data source for floor space prices in 2015 is the universe of housing transaction records provided by the Ministry of Land, Infrastructure, and Transport. Each record includes information on the location of a property (district), month and year of purchase, year built, lot size, etc. In order to obtain floor space prices representative for each district in 2015, I employ a Case-Shiller type repeated sales approach at the district level. To to so, I regress log of unit price on a set of dummies for year built, for month of purchase, and for year of purchase excluding 2015 along with district-level fixed effects. I use the estimated values of the district fixed effects (normalized such that the geometric mean is equal to 1) as my data for district-level floor space prices in 2015.

Additional District Level Characteristics

KOSIS (Korean Statistical Information System) provides a wide range of summary statistics describing district-level characteristics. I use the number of firms, number of firms discharging waster water, divorce rates, suicide rates, and geographical land area for each district to carry out cross-validation exercises comparing the model implied values of productivity and amenities with district-level characteristics. In addition, I collected information on the total land area used for residential purposes from the Land Use Statistics publicized by the Ministry of Land, Infrastructure, and Transport.

Annual Migration Rates

In order to understand the magnitude of migration rates across districts and across provinces (groups of districts), I leverage the restricted-access administrative data, which maintains the universe of migrant registry records in South Korea. This data is not used for the empirical

analysis of this paper because the records do not contain where migrants commute to. Notwith-standing its drawback, the records allow me to compute the annual migration rates and compare their magnitudes with the migration rates in the U.S.

B Supplementary Empirical Results

B.1 Local Income Taxes and Intergovernmental Transfers

Figure B.1: Spatial Distribution of Local Government Revenue by Sources
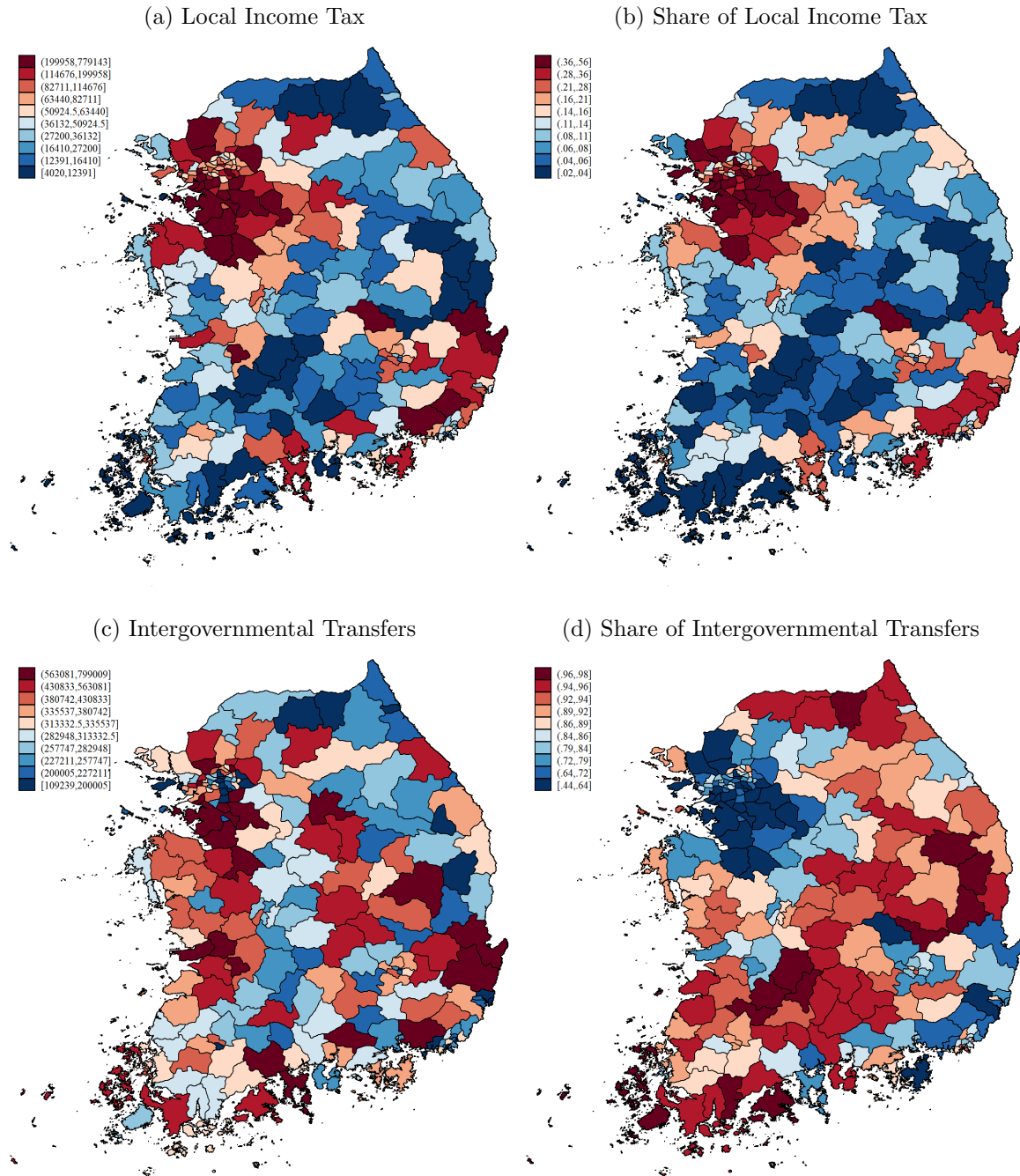
(a) Local Income Tax

(b) Share of Local Income Tax

(c) Intergovernmental Transfers

(d) Share of Intergovernmental Transfers

Notes: The figure on the left plots the spatial distribution of local government revenue by its sources (local income taxes and intergovernmental transfers) in 2015. The data source is the administrative data from the Ministry of Interior and Safety of South Korea. The denominator of the shares plotted in Panel (b) and (d) are the sum of local income tax and intergovernmental transfers for each district in 2015.

Figure B.2: Redistribution Paramters over Time

B.2 Decomposition of Observed Spatial Distribution of Workers

In order to understand the importance of the three spatial linkages (costs of migration, commuting, and job finding) in explaining the observed variation in the spatial distribution of workers $\pi_{orm,t}$, I carry out a type of variance decomposition exercise. First, I purge out location-specific factors $S_{o,t}$, $S_{r,t}$, and $S_{m,t}$ by residualizing $\pi_{orm,t}$ by the location specific fixed effects interacted with year dummies, $\phi_{o,t}$, $\phi_{r,t}$, and $\phi_{m,t}$. Second, I regress the residual on the fixed effects for each bilateral linkage, ϕ_{or}, ϕ_{rm}, and ϕ_{om} and obtain the predicted value $\hat{\pi}_{orm}$.[54] Mechanically, the predicted value is completely explained by ϕ_{or}, ϕ_{rm}, and ϕ_{om} together. By regressing $\hat{\pi}_{orm}$ on each of the pair-wise fixed effects of the bilateral linkages one at a time, I summarize how much

---

[54]The value of R^2 resulting from this regression is 0.72. Further interacting the set of bilateral fixed effects with year dummies only increases the value of R^2 to 0.79.

of the variation is explained by spatial linkages; the results are summarized in the table below. The migration linkage alone explains 41% of the variation; the commuting linkage explains 8%; the job finding alone explains 10%. Furthermore, an R^2 resulting from accounting any combination of two linkages together is higher than the sum of R^2's resulting from accounting each of the linkages separately.

Table B.1: Decomposition of Observed Variation in the Data

Regressors	R^2	Regressors	R^2
ϕ_{or}	0.4097	d_{or}	0.0700
ϕ_{rm}	0.0836	d_{rm}	0.0394
ϕ_{om}	0.1000	d_{om}	0.0563
ϕ_{or} and ϕ_{rm}	0.8590	ϕ_{or} and $\phi_{rm} + d_{om}$	0.9051
ϕ_{or} and ϕ_{om}	0.5498	ϕ_{or} and $\phi_{om} + d_{rm}$	0.8729
ϕ_{rm} and ϕ_{om}	0.3216	ϕ_{rm} and $\phi_{om} + d_{or}$	0.5154

Notes: This table reports the values of adjusted R^2 resulting from regressing the predicted spatial distribution of workers $\hat{\pi}_{r_0 rm}$ on the regressors listed in each row. The predicted distribution is computed based on the regression of the observed spatial distribution of workers residualized by location specific factors on the fixed effects of all three bilateral linkages of migration, commuting, and job finding.

Distances between origins and workplace locations together with the fixed effects of the migration and commuting linkages explain 91% of the observed variation. Distances of commuting together with the fixed effects of migration and job finding explain 87%. Lastly, distances of migration with the fixed effects of commuting and job finding accounts for 52% of the observed variations. Based on this simple exercise, I draw the following conclusions. First, net of the location specific factors, the spatial linkages of migration, commuting, and job finding are important determinants of the spatial distribution of workers. Second, the extent to which the commuting linkage explains the variation significantly improves along with the job finding linkage.

B.3 Inference

In this section, I discuss how I address issues related to estimating standard errors estimating the key elasticities of worker mobility. The concern overall is that the errors in each specification can be correlated in two ways. First, there is a classic clustering concern explained in Moulton (1990). Second, one may worry about the serial correlation over time within a panel dimension Bertrand et al. (2004). In order to address these concerns, I report standard errors that are robust to heteroskedasticity and allow multi-way clusterings.

First, with respect to estimating Equation (6), I allow errors to correlate across previous residences and across workplace locations sharing the same current residence in a given year. In addition, the serial correlation within each of the panel dimension (a triplet of previous residence, current residence, and workplace location) over time. Second, I conservatively cluster the standard errors at the migration-pair level when estimating Equation (14), at the commuting-pair level when estimating Equation (16), and at the job-finding-pair level when estimating Equation (18).

B.4   Travel Time vs. Distance of Commuting

Figure B.3: Travel Time vs. Distance of Commuting
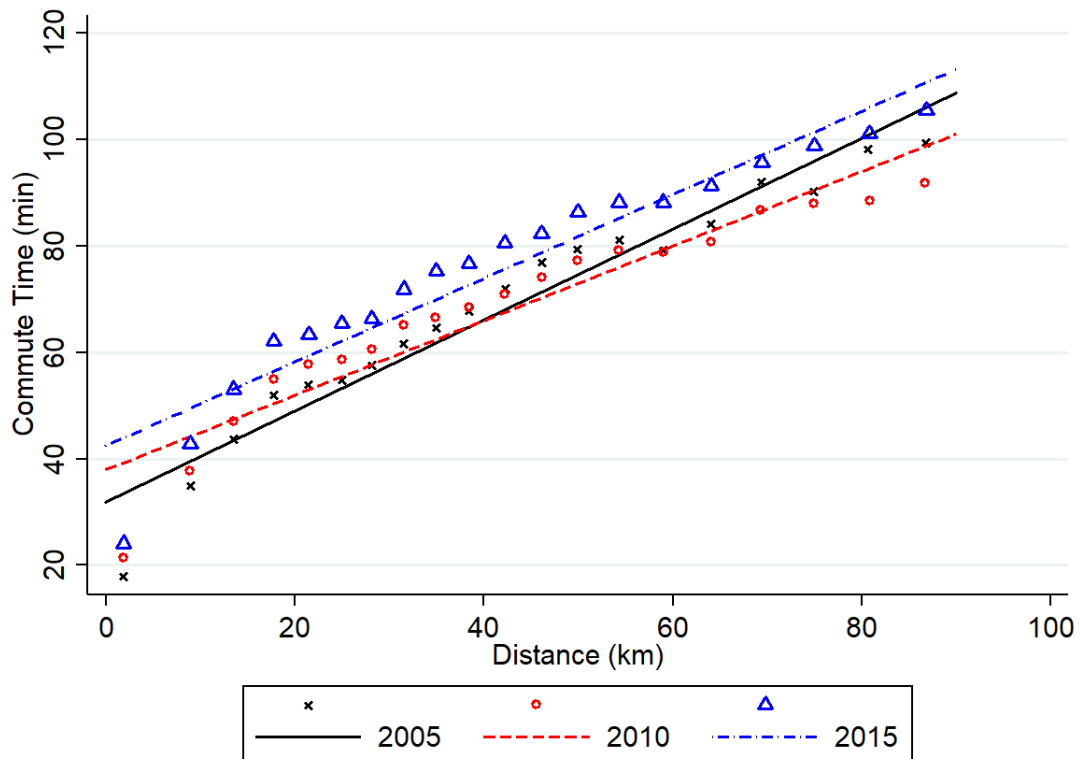
Notes: This figure plots average commuting time in minutes for each of 5 percentiles of commuting distance for each survey year (2005, 2010, and 2015) of the Population Census of South Korea.

As used in Ahlfeldt et al. (2015) and Morten and Oliveira (2018), an alternative measure to define the cost of commuting is travel time for commuting. Since travel time is surveyed in the Census, I compute average travel times in minutes for all bilateral commuting pairs. Figure B.3 shows a linear relationship between commuting distance and travel time. Furthermore, inspecting the relationship between commuting distance and time across 2005, 2010, and 2015,

71

Table B.2: Commuting Time (min) vs. Distance (km)

Dependent Variable:	(1)	(2)	(3)	(4)	(5)
	Commuting Time ($\tau_{rm,t}^{(time)}$)				
Distance (τ_{rm})	0.788***	0.915***	0.928***	1.017***	
	(0.00801)	(0.00784)	(0.00868)	(0.00839)	
$\tau_{rm} \times 2005$					1.066***
					(0.0121)
$\tau_{rm} \times 2010$					0.919***
					(0.0119)
$\tau_{rm} \times 2015$					1.061***
					(0.0121)
Observations	21,799	21,799	21,799	21,799	21,799
R^2	0.428	0.615	0.598	0.658	0.660
Fixed effects:					
Residence-Year ($\phi_{r,t}$)	N	Y	N	Y	Y
Workplace-Year ($\phi_{m,t}$)	N	N	Y	Y	Y

Notes: This table shows the relationship between distance of commuting and self-reported commuting time reported in the Population Census of South Korea. Each observation is a residence-workplace pair for each year of 2005, 2010, and 2015 with a positive number of workers reported to commute between residential and workplace locations. Robust standard errors in parentheses clustered at the residence-year, the workplace-year level: $***p < 0.01, **p < 0.05, *p < 0.1$.

there does not seem to be changes in commuting technology. To formalize, I estimate the following specification:

$$time_{rm,t} = \phi_{r,t} + \phi_{m,t} + \kappa^{time}d_{rm} + \varepsilon_{rm,t}^{time}.$$

The results are presented in B.2. Column (1) shows a raw correlation between distance and time. Across columns, I gradually introduce the fixed effects. According to Column (4), which corresponds to the equation above, travel time of commuting increases when distance of commuting increases by 1 kilometer. In order to understand whether or not this one-to-one relationship is stable over time, I re-estimate the equation above by interacting distance of commuting (time-invariant) with year dummies. The results are summarized in Column (5). The estimated coefficients for 2005, 2010, and 2015 are not statistically different from each other. I conclude that distance is a reasonable proxy for commuting time. The advantage of using travel times may be that measurement errors are averaged out by taking averages of travel times between localities observed at the individual-commuter level. However, average travel time changes over time, and such changes may be correlated with unobserved changes at the residence-workplace pair level that could also affect the spatial distribution of workers (e.g., an introduction of commuter rail). This is not the case for distances as they are fixed over time.

B.5 Omitted Variable Bias in OLS Estimates of Elasticities of Worker Mobility with respect to Local Government Goods and Home Prices

Table B.3: Elasticities of Worker Mobility with respect to Local Government Goods-OVB

Dependent Variable:	(1) $\ln \pi_{orm,t}$	(2) $\ln \pi_{orm,t}$	(3) $\ln \pi_{orm,t}$	(4) $\ln \pi_{orm,t}$	(5) $\ln \pi_{orm,t}$
Local Government Expenditure, $\ln G_{r,t}$ $(\beta_G = \lambda\epsilon)$	-0.231***	0.0965**	-0.433***	-0.452***	0.0957***
	(0.0149)	(0.0405)	(0.0213)	(0.0227)	(0.0299)
Number of Households, $\ln R_{r,t}$ $(\beta_R = \theta\lambda\epsilon)$	0.120***	0.0608**	0.480***	0.482***	0.590***
	(0.0125)	(0.0268)	(0.169)	(0.142)	(0.0522)
Floor Space Prices, $\ln Q_{r,t}$ $(\beta_Q = (1-\beta)\epsilon)$	-0.0416***	-0.0101	-0.0251*	-0.0431***	-0.00148
	(0.0129)	(0.0334)	(0.0135)	(0.0136)	(0.00653)
Observations	258,323	258,323	258,323	258,323	258,323
Fixed Effects:					
Job Finding Pairs $(\phi_{om,t})$	N	Y	N	N	Y
Migration Pairs (ϕ_{or})	N	N	Y	N	Y
Commuting Pairs (ϕ_{rm})	N	N	N	Y	Y

Notes: In this table, I report the OLS estimates of elasticities of worker mobility to local government expenditure, residential density, and home prices based on Equation 6, starting with a simple estimate without any fixed effects in Column (1). I introduce the fixed effects for job finding pairs interacted with time in Column (2), the fixed effects for migration pairs in Column (3), and adding the fixed effects for commuting pairs in Column (4). Column (5) reports the OLS estimates with all the fixed effects, separately introduced in Column (2)-(4) and corresponds to Equation 6. The sample is from 3 waves of the Population Census of South Korea in 2005, 2010, and 2015, based on 3,500,232 male household heads who are employed between the ages of 25 and 60. Each observation corresponds to a triplet of previous and current residences and workplace location. Robust standard errors in parentheses, with multi-way clustering by migration pair × year, commuting pair × year, and a triplet of previous and current residences and workplace: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

More formally, the directions of bias with respect to the OLS estimates are expressed as follows:

$$
\begin{bmatrix} \widehat{\beta}_G^{OLS} \\ \widehat{\beta}_R^{OLS} \\ \widehat{\beta}_R^{OLS} \end{bmatrix} = \begin{bmatrix} \beta_G \\ \beta_R \\ \beta_Q \end{bmatrix} + \frac{\begin{bmatrix} (\sigma_R^2\sigma_Q^2 - \sigma_{RQ}^2)\sigma_{G\zeta} + (\sigma_{GQ}\sigma_{RQ} - \sigma_{GR}\sigma_Q^2)\sigma_{R\zeta} + (\sigma_{GR}\sigma_{RQ} - \sigma_{GQ}\sigma_R^2)\sigma_{Q\zeta} \\ (\sigma_{GQ}\sigma_{RQ} - \sigma_{GR}\sigma_Q^2)\sigma_{G\zeta} + (\sigma_G^2\sigma_Q^2 - \sigma_{GQ}^2)\sigma_{R\zeta} + (\sigma_{GR}\sigma_{GQ} - \sigma_{RQ}\sigma_G^2)\sigma_{Q\zeta} \\ (\sigma_{GR}\sigma_{RQ} - \sigma_{GQ}\sigma_R^2)\sigma_{G\zeta} + (\sigma_{GR}\sigma_{GQ} - \sigma_{RQ}\sigma_G^2)\sigma_{R\zeta} + (\sigma_G^2\sigma_R^2 - \sigma_{GR}^2)\sigma_{Q\zeta} \end{bmatrix}}{\sigma_G^2\sigma_R^2\sigma_Q^2 + 2\sigma_{GR}\sigma_{RQ}\sigma_{GQ} - (\sigma_G^2\sigma_{RQ}^2 + \sigma_R^2\sigma_{GQ}^2 + \sigma_Q^2\sigma_{GR}^2)}
$$

(38)

where $\sigma_G^2\sigma_R^2\sigma_Q^2 + 2\sigma_{GR}\sigma_{RQ}\sigma_{GQ} - (\sigma_G^2\sigma_{RQ}^2 + \sigma_R^2\sigma_{GQ}^2 + \sigma_Q^2\sigma_{GR}^2) > 0$ by the Cauchy-Schwarz inequality. Also, note that all the variance and covariance terms are conditional on the set of fixed effects (fixed effects of origin-workplace-by-year, migration pair, commuting pair).

B.6 Validity of Instrumental Variables based on the Tax Reforms with Sorting

The quantitative spatial model I present in this paper assumes that the workers are born with initial residences and have heterogeneous preferences for locations. They are otherwise homogeneous. Therefore, I do not take a stance in potential reallocation of workers based on sorting. However, a residence with a greater share of its residents with higher education (skill) may generate a higher amenity value relative to other residences (Diamond, 2016). In this case, the error term in Equation (6) would include the distribution of workers by education $\pi_{edu|r,t}$.

The exclusion restriction (8) is violated due to sorting only if the fiscal reforms resulted in making residences relatively more or less attractive by changing the educational composition within districts. Since I observe the contemporaneous shares of workers by education levels from the Population Census of South Korea, I can test whether the tax reforms directly affected the educational composition of workers at their residences. I consider the following specification:

$$\pi_{b|r,t} = \phi_r + \eta_{b',b}\tau_{b',t} + \zeta_{b,r,t}, \tag{39}$$

where the dependent variable $\pi_{b|r,t}$ is the demeaned fraction of workers with educational level b (low and high, which proxy the low and high income brackets in the tax schedule) living in residence r in year t; the residence fixed effects ϕ_r captures the baseline differences in the dependent variable; $\tau_{b',t}$ is the tax rates in year t for income bracket b'. With the residence fixed effects, if an estimated value of $\eta_{b',b}$ is statistically different from zero, then I reject the hypothesis that the changes in tax rates for income bracket b' had no impact on the changes in the distribution of workers with education level b.

Table B.4 reports the estimation results. All the coefficients are not statistically different from zero, nor are they economically significant. In sum, I draw a conclusion that the tax reforms did not result in changes in the attractiveness of residences based on their educational composition of workers. Therefore, predicted tax contributions by low and high income groups are orthogonal to the contemporaneous education distribution.

B.7 Alternative (Parsimonious) Specification to Estimate the Elasticities of Worker Mobility with respect to Local Government Goods and Home Prices

$$\ln \pi_{orm,t} = \underbrace{\ln \phi_{o,t}}_{\tilde{\phi}_{o,t}} + \underbrace{\ln \phi_{m,t}}_{\tilde{\phi}_{m,t}} + \ln(\varepsilon_{or,t}\varepsilon_{rm,t}\varepsilon_{om,t}\varepsilon'_{orm,t}D_{or}D_{rm}D_{om})^{-\epsilon} + \beta_G \ln G_{r,t} - \beta_R \ln R_{r,t} - \beta_Q \ln Q_{r,t} + \ln \tilde{B}_{r,t}$$

$$\tag{40}$$

Equation 40 is the expression for the log transformation of the gravity equation (5), augmented

Table B.4: Tax reforms did not affect education distribution

	(1)	(2)
A. Educational Attainment: Low $(\pi_{low|r,t})$		
Tax Rate (Low) $\tau_{low,t}$	3.79e-09	
	(0.00115)	
Tax Rate (High) $\tau_{high,t}$		2.38e-09
		(0.000738)
B. Educational Attainment: High $(\pi_{high|r,t})$		
Tax Rate (Low) $\tau_{low,t}$	-1.35e-09	
	(0.000582)	
Tax Rate (High) $\tau_{high,t}$		-6.44e-10
		(0.000382)
Observations	666	666

Notes: This table reports the estimation results based on Equation (39). Each estimated coefficient corresponds to the effect of changes in tax rates on changes in the educational composition of residences. The sample is constructed from 3 waves of the Population Census of South Korea in 2005, 2010, and 2015, based on 3,494,198 individual household heads who are employed between the ages of 25 and 60. Each observation corresponds to a residence for each year. Robust standard errors in parentheses clustered at the residence level: $***p < 0.01, **p < 0.05, *p < 0.1$.

with the time subscripts wherever applicable. For expository purposes, I unpack the stochastic error term $\varepsilon_{orm,t}$ into four components: $\varepsilon_{orm,t} = \varepsilon_{or,t}\varepsilon_{rm,t}\varepsilon_{om,t}\varepsilon'_{orm,t}$; I assume each of $\varepsilon_{om,t}$ and $\varepsilon'_{orm,t}$ follows a log normal distribution with mean equal to 1. I consider a parsimonious specification alternative to the main estimating equation (6) as follows:

$$\ln \pi_{orm,t} = \underbrace{\ln \phi_{o,t}}_{\tilde{\phi}_{o,t}} + \underbrace{\ln \phi_{m,t}}_{\tilde{\phi}_{m,t}} + \phi_{or} + \phi_{rm} + \beta_G \ln G_{r,t} - \beta_R \ln R_{r,t} - \beta_Q \ln Q_{r,t} + \underbrace{\ln \tilde{B}_{r,t}\varepsilon_{om,t}\varepsilon'_{orm,t}}_{\zeta'_{orm,t}}. \quad (41)$$

The difference between Equation 5 and Equation 41 is that the stochastic error term $\varepsilon_{om,t}$ is loaded onto the error term $\zeta'_{orm,t}$ in Equation 41. Both specifications are consistent with the model. I summarize the OLS estimates, first-stage estimates, and 2SLS estimates based on Specification 41 in Table B.5. The results are qualitatively and quantitatively similar to the results reported in Table 3.

Table B.5: (2SLS) Parsimonious FE

Dependent Variable:	(1) OLS $\ln \pi_{orm,t}$	(2) First Stage $\ln G_{r,t}$	(3) First Stage $\ln R_{r,t}$	(4) First Stage $\ln Q_{r,t}$	(5) 2SLS $\ln \pi_{orm,t}$
Local Government Expenditure, $\ln G_{r,t}$ ($\beta_G = \lambda\epsilon$)	0.106***				1.105***
	(0.016)				(0.190)
Number of Households, $\ln R_{r,t}$ ($-\beta_R = -\theta\lambda\epsilon$)	0.565***				-0.807***
	(0.023)				(0.296)
Floor Space Prices, $\ln Q_{r,t}$ ($-\beta_Q = -(1-\beta)\epsilon$)	-0.0085				-0.528***
	(0.0039)				(0.039)
Predicted Tax Contribution (low), $IV_{r,t}^{low}$		13.82***	7.054***	48.62***	
		(0.210)	(0.152)	(0.901)	
Predicted Tax Contribution (high), $IV_{r,t}^{high}$		13.86***	6.990***	18.86***	
		(0.099)	(0.071)	(0.422)	
Number of Households 30 years ago, $IV_{r,t}^{R}$		0.028***	-0.018***	0.110***	
		(0.001)	(0.001)	(0.005)	
Observations	258,323	258,323	258,323	258,323	258,323
SW F-stat		-	-	-	
$\hat{\theta}$	-6.164**				0.731**
	(2.176)				(0.152)

Notes: In this table, I compare the OLS estimates and 2SLS estimates of elasticities of worker's mobility to local government expenditure and resident population levels based on Equation 41. Column (1) reports the OLS estimates. Column (2) and Column (3) report the first stage results. The 2SLS estimates are reported in Column (4). Across columns, I replace the pairwise fixed effects of job finding interacted with year dummies in the main specification (6) with a more parsimonious set of fixed effects for previous residence by year, workplace location by year, and job finding pairs. The sample ($N = 258,323$) is from 3 waves of the Population Census of South Korea in 2005, 2010, and 2015, based on 3,500,232 male household heads who are employed between the ages of 25 and 60. Each observation corresponds to a triplet of previous and current residences and workplace location. Robust standard errors for Column (1), (2), and (3) and bootstrapped (20,000 replications) standard errors for Column (4) in parentheses, with multi-way clustering by migration pair × year, commuting pair × year, and a triplet of previous and current residences and workplace: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

B.8 2SLS Estimation based on Migration and Commuting Flows

Recall the gravity equation of the model is given by:

$$\pi_{orm,t} = \frac{(\tilde{B}_{r,t}(1-\tau_{m,t})\tilde{w}_{m,t}G_{r,t}^\lambda)^\epsilon \pi_{o,t}}{(\varepsilon_{orm,t}D_{or}D_{rm}D_{om}Q_{r,t}^{1-\beta}R_{r,t}^{\theta\lambda})^\epsilon} \Bigg/ \underbrace{\sum_{r'=1}^{J}\sum_{m'=1}^{J}\frac{(\tilde{B}_{r',t}(1-\tau_{m',t})\tilde{w}_{m',t}G_{r',t}^\lambda)^\epsilon}{(\varepsilon_{or'm',t}D_{or'}D_{r'm'}D_{om}Q_{r',t}^{1-\beta}R_{r',t}^{\theta\lambda})^\epsilon}}_{\Phi_{o,t}}$$

Summing it over workplace location, I derive an expression for migration flow:

$$\pi_{om,t} = \frac{(\tilde{B}_{r,t}G_{r,t}^\lambda)^\epsilon \pi_{o,t}/\Phi_{o,t}}{(D_{or}Q_{r,t}^{1-\beta}R_{r,t}^{\theta\lambda})^\epsilon}\sum_{m=1}^{J}\left(\frac{(1-\tau_{m,t})\tilde{w}_{m,t}}{\varepsilon_{orm,t}D_{rm}D_{om}}\right)^\epsilon$$

Then, I derive an estimating equation by taking the log transformation:

$$\ln \pi_{or,t} = \phi_{o,t} + \phi_{or} + \underbrace{\lambda\epsilon}_{\beta_G}\ln G_{r,t} - \underbrace{\theta\lambda\epsilon}_{\beta_R}\ln R_{r,t} - \underbrace{(1-\beta)\epsilon}_{\beta_Q}\ln Q_{r,t} + \zeta_{or,t}^{mig}, \qquad (42)$$

where $\phi_{o,t} = \ln \phi_{o,t}/\Phi_{o,t}$, $\phi_{or} = \ln D_{or}^\epsilon$, capturing the cost of migration; $\zeta_{or,t}^{mig} = \ln \tilde{B}_{r,t}\underbrace{\sum_{m=1}^{J}\left(\frac{(1-\tau_{m,t})\tilde{w}_{m,t}}{\varepsilon_{orm,t}D_{rm}D_{om}}\right)^\epsilon}_{ALMA_{or,t}}$.

The relevance of the instrumental variables (tax reforms and historical residential density) holds as when using the worker mobility (migration and commuting jointly). The exclusion restriction requires:

$$E\begin{bmatrix} IV_{r,t}^{low}\zeta_{or,t}^{mig} & |\phi_{o,t},\ \phi_{or} \\ IV_{r,t}^{high}\zeta_{or,t}^{mig} & |\phi_{o,t},\ \phi_{or} \\ IV_{r,t}^{R}\zeta_{or,t}^{mig} & |\phi_{o,t},\ \phi_{or} \end{bmatrix} = 0.$$

This is violated because $IV_{r,t}^{low}$ and $IV_{r,t}^{high}$ are functions of tax rates. And, the error terms includes $ALMA_{or,t}$, which is also a function of tax rates. Therefore, 2SLS estimates would be inconsistent. In particular, the direction of bias in 2SLS estimate for β_G is downward $\because cov(G_{r,t}, \zeta_{or,t}^{mig} < 0)$. Note that if D_{rm} is equal to 1 (i.e., no spatial friction from commuting), then exclusion restriction is satisfied. Therefore, observing biases in 2SLS estimates using migration implies it is important to take commuting into account.

Similarly, I sum $\pi_{orm,t}$ over previous residence, I derive an expression for commuting flow:

$$\pi_{rm,t} = \frac{(\tilde{B}_{r,t}(1-\tau_{m,t})\tilde{w}_{m,t}G_{r,t}^\lambda)^\epsilon}{(D_{rm}Q_{r,t}^{1-\beta}R_{r,t}^{\theta\lambda})^\epsilon}\sum_{o=1}^{J}\frac{\pi_{o,t}/\Phi_{o,t}}{(\varepsilon_{orm,t}D_{or}D_{om})^\epsilon}$$

Then, I derive an estimating equation by taking the log transformation as follows:

Table B.6: Comparison of Estimation Results based on Migration and Commuting Flows

	(1) Both	(2) Both	(3) Migration Flows	(4) Migration Flows	(5) Commuting Flows	(6) Commuting Flows
	OLS	2SLS	OLS	2SLS	OLS	2SLS
VARIABLES	$\ln \pi_{orm,t}$	$\ln \pi_{orm,t}$	$\ln \pi_{or,t}$	$\ln \pi_{or,t}$	$\ln \pi_{rm,t}$	$\ln \pi_{rm,t}$
Local Government Expenditure, $\ln G_{r,t}$ ($\beta_G = \lambda\epsilon$)	0.0957***	1.072***	0.357***	-1.522***	0.302***	4.935***
	(0.0299)	(0.387)	(0.0197)	(0.188)	(0.0315)	(0.749)
Number of Households, $\ln R_{r,t}$ ($\beta_R = \theta\lambda\epsilon$)	0.590***	-0.844	1.118***	3.205***	1.113***	-3.293***
	(0.0522)	(0.622)	(0.0361)	(0.267)	(0.0512)	(0.852)
Floor Space Prices, $\ln Q_{r,t}$ ($\beta_Q = (1-\beta)\epsilon$)	-0.00148	-0.490***	0.0424***	0.568***	-0.0288***	-2.011***
	(0.00653)	(0.0672)	(0.00386)	(0.0226)	(0.00729)	(0.222)
Observations	258,323	258,323	70,427	70,427	20,676	20,676
Fixed Effects:						
$\phi_{om,t}, \phi_{or}, \phi_{rm}$	Y	Y	N	N	N	N
$\phi_{o,t}, \phi_{or}$	N	N	Y	Y	N	N
$\phi_{m,t}, \phi_{rm}$	N	N	N	N	Y	Y

Notes: In Column (1) and (2), I report the OLS and 2SLS estimates of the effects of local government spending, residential density, and housing prices based on worker mobility defined in terms of both migration and commuting. In Column (3) and (4), I report the OLS and 2SLS estimates based on migration flows alone. In Column (5) and (6), I report the OLS and 2SLS estimates on commuting flows alone.

$$\ln \pi_{rm,t} = \phi_{m,t} + \phi_{rm} + \underbrace{\lambda\epsilon}_{\beta_G} \ln G_{r,t} - \underbrace{\theta\lambda\epsilon}_{\beta_R} \ln R_{r,t} - \underbrace{(1-\beta)\epsilon}_{\beta_Q} \ln Q_{r,t} + \zeta^{com}_{rm,t}, \qquad (43)$$

where $\phi_{m,t} = \ln((1-\tau_{m,t})\tilde{w}_{m,t})^{\epsilon}$; $\phi_{rm} = \ln D^{\epsilon}_{rm}$ which captures the cost of commuting; $\zeta^{com}_{rm,t} =$
$\ln \tilde{B}_{r,t} \underbrace{\sum_{o=1}^{J} \frac{\pi_{o,t}/\Phi_{o,t}}{(\varepsilon_{orm,t} D_{or} D_{om})^{\epsilon}}}_{AMMA_{rm,t}}$. Similarly, the relevance of the instrumental variables (tax reforms
and historical residential density) still holds since the relevance did not hinge on the assumptions
of spatial frictions. The exclusion restriction requires,

$$E \left[\begin{array}{cc} IV^{low}_{r,t}\zeta^{com}_{or,t} & |\phi_{m,t},\ \phi_{rm} \\ IV^{high}_{r,t}\zeta^{com}_{or,t} & |\phi_{m,t},\ \phi_{rm} \\ IV^{R}_{r,t}\zeta^{com}_{or,t} & |\phi_{m,t},\ \phi_{rm} \end{array} \right] = 0.$$

The exclusion restriction in this case is violated. In this case, $IV^{low}_{r,t}$ and $IV^{high}_{r,t}$ are functions
of tax rate, while the error terms includes $AMMA_{or,t}$ also a function of tax rates because $\Phi_{o,t}$
includes tax rates. Therefore, 2SLS estimates would be inconsistent. In particular, the direction
of bias in 2SLS estimate for β_G is upward $\because cov(G_{r,t}, \zeta^{com}_{rm,t} > 0)$. Note that if D_{or} is equal to 1
(no spatial friction from migration), then exclusion restriction is satisfied. Therefore, observing
biases in 2SLS estimates based on commuting flows implies that migration needs to be taken
into account.

In Table B.6, I report the OLS and 2SLS estimates based on Equation (6) in Column (1)
and (2) using both migration and commuting flows, Equation (42) in Column (3) and (4) using
migration flows alone, and Equation (43) using commuting flows alone. The results altogether
show that in order to consistently estimate the elasticities of interest leveraging the tax reforms,
both migration and commuting need to be considered jointly.

($\phi_{m,t}$-OVB) If a district is located near employment locations with high wages, then the
average income of the residents in this district is high. Because of the redistributive intergov-
ernmental transfers, the local government expenditure is low. The fixed effects for employment
locations address the omitted variable bias rising from the negative correlation between local
labor market returns and local government spending.

($\phi_{m,t}$-Exclusion Restriction) Without the fixed effects for workplace location, the exclusion
restriction of the proposed instruments based on tax reforms is violated because the tax rates
directly affects worker mobility.

($\phi_{o,t}$-OVB) There are two factors specific to origin: the initial distribution of workers across
residential location and multilateral resistance. So, the fixed effects capture the effects of
augmented migrant market access. If a residence is situated around places with higher values of

migrant market access, this residence is likely to have a greater number of migrants, resulting in a higher residential density. In turn, higher population is likely to be positively correlated with local spending because of a high tax base and redistribution favoring dense localities.

($\phi_{o,t}$-Exclusion Restriction) Exclusion redistribution is violated unless these fixed effects are introduced because the multilateral resistance term is a function of tax rates.

C   Supplementary Quantitative Results

C.1   Adjusted After-Tax Wages and Fréchet Shape Parameter

$$\pi_{m|rr_0} = \frac{M_m \left(\frac{(1-\tau_m)w_m}{D_{r_0 m} D_{rm}} \right)^{\epsilon}}{\sum_{m'=1}^{S} M_{m'} \left(\frac{(1-\tau_{m'})w_{m'}}{D_{r_0 m'} D_{rm'}} \right)^{\epsilon}} \tag{44}$$

Take log transformation both sides. Then, add the costs of job finding and commuting. Then, regress the left hand side variable on the fixed effects of workplace location and migration pairs. I recover the adjusted after-tax wages from the estimated fixed effects of workplace locations.

I estimate the dispersion parameter by taking the ratio of the dispersion of adjusted after-tax wages and the dispersion of observed after-tax wages. The estimated value of ϵ is equal to 3.5, which is statistically significantly different from zero at the 1 percent.
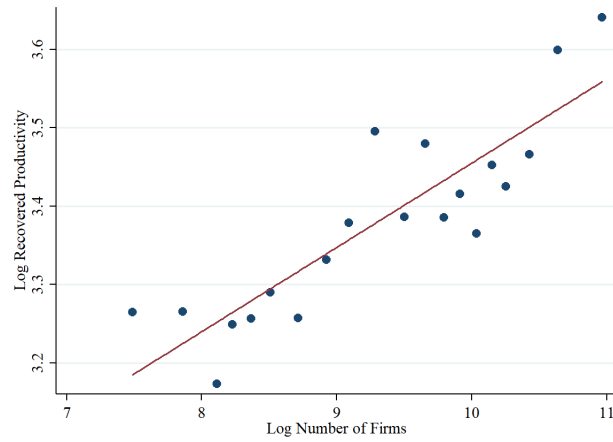
C.2 Local Productivity
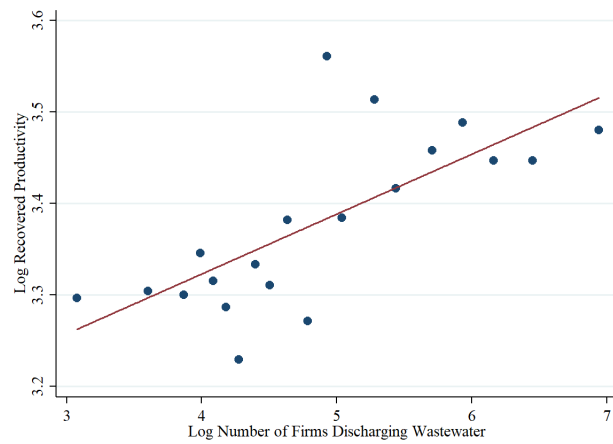
Figure C.1: Spatial Distribution of Productivities
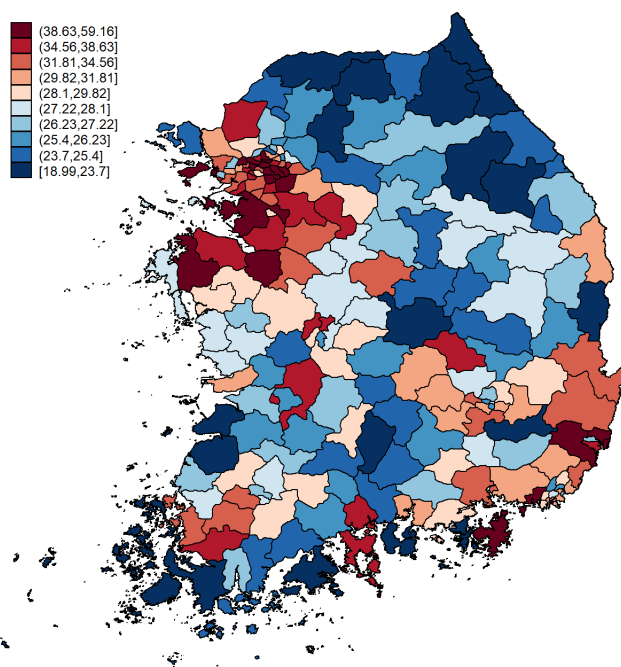
Notes: This figure plots the recovered values of productivity for each district using the model with the data in 2015. Section 7.3 explains how the values are recovered from the estimated fixed effects.

Figure C.2: Recovered Local Amenities vs. Number of Firms
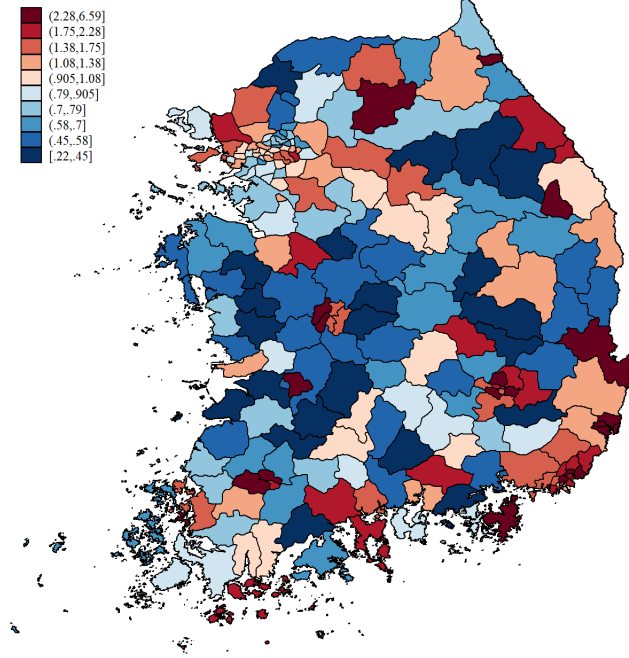
(a) Number of Firms

(b) Number of Dirty Firms

Notes: This figure plots the values of log productivity recovered in Section 7.3 against number of firms in Panel (a) and firms discharging wastewater in Panel (b) in 2015. Each point corresponds to 5 percentile of number of firms.

C.3 Adjusted Amenities

Figure C.3: Recovered Amenities in 2015

Notes: This figure plots the recovered amenity values for each district using the model with the data in 2015. Section 7.3 explains how the amenity values are recovered from the estimated fixed effects.

C.4 Fiscal Decentralization Policy Parameters

- Observed Data: Total Expenditure G_r and its sources: local tax revenue LT_r and inter-governmental transfers IT_r

- Local government spending:

$$G_r = \varsigma \sum_{m=1}^{S} \tau_m w_m \pi_{m|r} R_r + \varsigma_r (1-\varsigma) \chi \sum_{r'=1}^{S} \sum_{m'=1}^{S} \tau_{m'} w_{m'} \pi_{m'|r'} R_{r'} \qquad (45)$$

where $\tilde{\varsigma}$ denotes the fraction of total local tax revenue delivered to the national government $1 - \varsigma = 0.9$ multiplied by the fraction of the national tax revenue used for redistribution $\chi = 0.35$. This implies that $(1 - \chi)(1 - \varsigma) = 0.5915$ of the local tax revenue is used for the national government. This also means that in total about 40 percent of the local tax revenue (i.e., extent of fiscal decentralization) is spendt locally. When I conduct counterfactual policy experiments. I keep the extent of fiscal decentralization constant at 40 percent and only change the extent of redistribution, ranging from 0 to 40 percent. Also, I keep the rules of redistribution ($\{\varsigma_j\}_{j=1}^{J}$, where $\varsigma_j = \frac{IT_j}{\sum_{j'=1}^{J} IT_j}$) constant at the

Figure C.4: Recovered Local Amenities vs. Measures of Quality of Life

(a) Suicide Rate
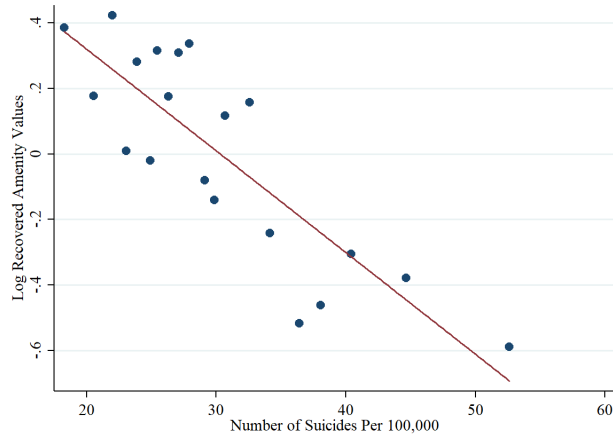
(b) Divorce Rate
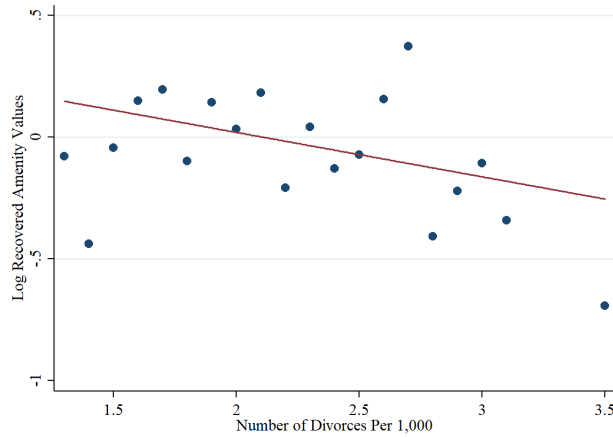
Notes: This figure plots the values of adjusted amenities recovered in Section 7.3 against two measures proxying the quality of life observed in 2015: suicides per 100,000 residents in Panel (a); number of divorces per 1,000 couple in Panel (b). Each point corresponds to 5 percentile of the quality-of-life measures.

2015 values. When the extent of redistribution is equal to 0%, local government spending is solely financed by local tax revenue from residents. When it is equal to 40%, local government spending is completely determined by intergovernmental transfers.

Figure C.5: Redistribution Parameters in 2015
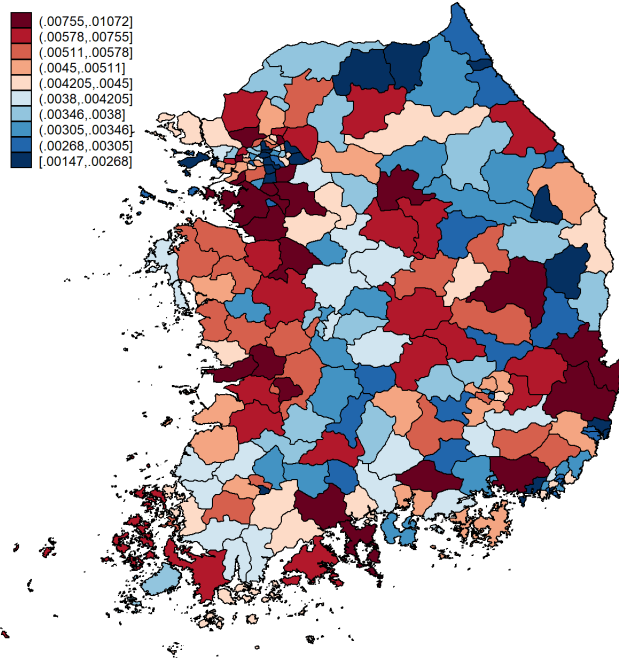
Notes: This figure plots the observed values of rules of redistribution (ς_r) in 2015 for each residence. The districts in red receives greater shares of intergovernmental transfers from the national government than the districts in blue. The shares sum up to 1.

C.5 Algorithm to Solve the Model

I briefly describe the iterative algorithm used to solve for the equilibrium of the model (Ahlfeldt et al., 2015; Monte et al., 2018; Tsivanidis, 2019); See Appendix C.5 for details. Section 6.5 characterizes the equilibrium of the model and the system of equations to be solved. First, I make initial guess for a set of endogenous variables. Second, using these initial values, I solve the system of equations of the model for a new value of the endogenous variables. Third, I update the guess for the equilibrium by taking a weighted average of the initial and the new values. Lastly, I iterate this process until the new and initial values converge.

I solve for stock of floor space for each district appealing to the market clearing for floor space in Section 6.3. First, in equilibrium, the residential floor space demanded is a function of after-tax wages ($(1-\tau)w_m$), conditional commuting probabilities ($\pi_{m|r}$), residential population (R_r), and per-unit floor space prices (Q_r) given housing expenditure share ($1-\beta$) as in Equation (27). Second, the commercial floor space demanded is determined by local productivity (A_j),

Table C.1: Goodness of Fits: Model vs. Alternatives (Migration)

$\frac{\sum(X_r^{restriction}-X_r^{data})^2}{\sum(X_r^{baseline}-X_r^{data})^2}$	Restrictions					
	(1)	(2)	(3)	(4)	(5)	(6)
	$\rho = 0$	$\delta = 0$	$\rho = 0$ $\delta = 0$	$\rho = 0$ $\delta = 0$ $\kappa = \kappa^{lit}$	$\kappa \to \infty$	$\rho = \rho^{lit}$ $\delta = 0$ $\kappa \to \infty$
$X_r =$						
R_r	1.16	1.09	2.47	2.61	1.34	3.60
G_r	1.12	1.10	1.87	1.86	1.95	2.83
L_m	1.19	1.14	3.16	3.17	1.97	4.81

Notes: In this table, I show the goodness of fits under alternative assumptions on the spatial frictions relative to the baseline model. Each value reported in this table corresponds to the sum of squared residuals relative to the baseline model. Therefore, a value higher than 1 implies that the base line model performs better in predicting the observed values. I solve for an equilibrium assuming that migration is costless in Column (1), job finding is costless in Column (2), both are costless in Column (3), and both are costless with the spatial decay of commuting equal to the value estimated following the commuting literature (Column (5) of Table 5). In Column (5), I assume that commuting is prohibitively costly. Column (6) assumes the value of spatial decay of migration equal to the value estimated following the migration literature (Column (5) of Table 4) while job finding is costless.

employment population (L_j), and floor space prices (Q_j) given labor share in production (α) as in Equation (28). I set floor space stock of a district equal to the sum of floor space demands for residential and commercial uses computed based on the tax rate from Section 7.2 and local productivity recovered above as well as the observed data on wages, floor space prices, conditional commuting probabilities, and residential and employment population.

C.6 Goodness of Fits relative to Alternative Specifications

D Supplementary Theoretical Results

D.1 Derivation of the Gravity Equation in Section 3

Because the indirect utility is equal to the idiosyncratic component of utility (z_{irm}) multiplied by the indirect utility of the systematic component (v_{orm} in Equation (3)), the distribution of utility for a worker from origin o living in district r and working in district m is also Fréchet distributed. Therefore, the cumulative distribution function of the utility is

$$F_{rm}(u) = \Pr[U \leq u] = \Pr(z \leq u \times v_{orm}^{-1}), \tag{46}$$

where $z \sim G(z) = \exp(-T_r M_m z^{-\epsilon})$. It follows that

$$F_{rm}(u) = \exp\left(-\frac{T_r M_m B_r (1-\tau_m) w_m}{D_{orm} Q_r^{1-\beta}} \left(\frac{G_r}{R_r^\theta}\right)^\lambda u^{-\epsilon}\right) \equiv \exp(-\Phi_{orm} u^{-\epsilon}). \tag{47}$$

I denote f_{rm} to be the density function. Conditional on their origin o, workers choose a pair of residence r and workplace m that achieves that maximum utility. Therefore, the probability of choosing a residence-workplace pair (residence r and workplace location m) conditional on having come from origin o is expressed as follows:

$$\pi_{rm|o} = \Pr[u_{rm|o} \geq \max\{u_{jk}\}; \forall j, k]$$

$$= \int_0^\infty \prod_{k \neq j} F_{rk}(u) \times \left(\prod_{j \neq r} \prod_k F_{jk}(u) \right) f_{rm}(u) du$$

$$= \int_0^\infty \prod_j \prod_k \epsilon \Phi_{orm} u^{-(\epsilon+1)} \exp(-\Phi_{ojk} u^{-\epsilon}) du$$

$$= \int_0^\infty \epsilon \Phi_{orm} u^{-(\epsilon+1)} \exp(-\Phi_o u^{-\epsilon}) du,$$

where $\Phi_o = \sum_{r=1}^J \sum_{m=1}^J \Phi_{orm}$. Evaluating the integral above, the probability of choosing residence r and workplace m conditional on origin o is:

$$\pi_{rm|o} = \frac{T_r M_m \left(\frac{B_r (1-\tau_m) w_m}{D_{orm} Q_r^{1-\beta}} \left(\frac{G_r}{R_r^\theta} \right)^\lambda \right)^\epsilon}{\sum_{r'=1}^J \sum_{m'=1}^J T_{r'} M_{m'} \left(\frac{B_{r'} (1-\tau_{m'}) w_{m'}}{D_{or'm'} Q_{r'}^{1-\beta}} \left(\frac{G_{r'}}{R_{r'}^\theta} \right)^\lambda \right)^\epsilon} \equiv \frac{\Phi_{orm}}{\Phi_o} \tag{48}$$

Because the maximum of a sequence of Fréchet distributed random variables is itself Fréchet distributed. Therefore,

$$F_o(u) = \exp(-\Phi_o u^{-\epsilon}), \text{ where } \Phi_o = \sum_{r=1}^J \sum_{m=1}^J \frac{T_r M_m B_r (1-\tau_m) w_m}{D_{orm} Q_r^{1-\beta}} \left(\frac{G_r}{R_r^\theta} \right)^\lambda. \tag{49}$$

Based on the distribution of utility defined above, the expected utility for workers with origin o is given by:

$$\mathbb{E}[u|o] = \int_0^\infty \epsilon \Phi_o u^{-\epsilon} e^{-\Phi_o u^{-\epsilon}} du = \Gamma \left(\frac{\epsilon-1}{\epsilon} \right) \Phi_o^{1/\epsilon} \equiv \bar{u}_o. \tag{50}$$

D.2 Isomorphism of the Gravity Equation

I show that the gravity equation (5) is isomorphic to the types of gravity equations derived in the literature on costly movements of people: commuting and migration.

D.2.1 Commuting Literature

The literature on commuting decisions assume free mobility in terms of migration. Therefore, there is usually no discussion on how workers are distributed across space before they make their commuting decisions. The underlying assumption in this literature is that there is no cost of enabling each commuting possibility via migration and job finding. This assumption translate to setting both ρ and δ equal to zero in my model presented in Section 3. Then, the distribution of workers by current residence and workplace is independent to the distribution of workers by initial residence. Therefore, Equation (5) does not vary by initial residence o and is given by:

$$\pi_{rm} = \frac{T_r M_m \left(\frac{B_r(1-\tau_m)w_m}{D_{rm}Q_r^{1-\beta}} \left(\frac{G_r}{R_r^\theta} \right)^\lambda \right)^\epsilon}{\sum_{r'=1}^{S} \sum_{m'=1}^{S} T_{r'} M_{m'} \left(\frac{B_{r'}(1-\tau_{m'})w_{m'}}{D_{r'm'}Q_{r'}^{1-\beta}} \left(\frac{G_{r'}}{R_{r'}^\theta} \right)^\lambda \right)^\epsilon}, \tag{51}$$

where D_{rm} is a commuting cost, a function increasing in distance between r and m. Further assuming no tax on wage (i.e., $\tau_m = 0$ for all m) and no utility derived from local government goods and services (i.e., $\lambda = 0$), Equation (51) is identical to the gravity equations based on the spatial models of Ahlfeldt et al. (2015) and Monte et al. (2018).

D.2.2 Migration Literature

The literature on migration decisions generally considers movements of people across relatively larger spatial units such that workers are likely to work and live in the same spatial unit upon migrating. Accordingly, in this literature, there is no distinction between a workplace and a residence since workers are assumed to work and live in the same locations. This assumption can be implemented in my model by setting the commuting cost to a workplace outside of residence equal to ∞. Then, the migration patterns of workers are summarized by:

$$\pi_{or} = \frac{T_r M_r \left(\frac{B_r(1-\tau_r)w_r}{D_{or}Q_r^{1-\beta}} \left(\frac{G_r}{R_r^\theta} \right)^\lambda \right)^\epsilon \pi_o}{\sum_{r'=1}^{S} T_{r'} M_{r'} \left(\frac{B_{r'}(1-\tau_{r'})w_{r'}}{D_{or'}Q_{r'}^{1-\beta}} \left(\frac{G_{r'}}{R_{r'}^\theta} \right)^\lambda \right)^\epsilon}, \tag{52}$$

where D_{or} is the iceberg cost associated with migration. Again, assuming to tax on wage and no benefits from local government goods and services, Equation (52) shares the same structure as the gravity equations based on the spatial models of migration considered in Bryan and Morten (2018) and Morten and Oliveira (2018).