The Spatial Distribution of Public Spending, Commuting, and Migration*

Wookun Kim†

April 12, 2023

Click Here for the Most Recent Version

Abstract

What are the aggregate welfare consequences of fiscal transfers across local governments that finance their spending? Answering this question requires an understanding of how much people value local public spending. I develop a spatial equilibrium framework in which workers' simultaneous migration and commuting choices reveal preferences. I combine this framework with unique data from South Korea and leverage tax reforms as a source of exogenous variation. The estimated mobility responses imply that workers value an additional dollar of per-capita local government spending at 75 cents of their after-tax income. 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 jointly made. Ignoring either 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: local public finance, redistribution, gravity equation, migration, commuting, quantitative spatial model

JEL codes: H3, H77, J61, R12, R13, R5

*I am extremely grateful to my advisor Pablo Fajgelbaum for his guidance and support. I thank Adriana Lleras-Muney, Kathleen McGarry, and Jonathan Vogel for their encouragement and suggestions. Youssef Benzarti and Manisha Shah provided valuable advice. I acknowledge Klaus Desmet, Gabriel Ahlfeldt, Brett McCully, and many other participants at Korea University, Sogang University, Yonsei University, HKUST, GRIPS, LKYSPP NUS, Federal Reserve Bank of Atlanta, Korea Institute of Public Finance, Southern Methodist University, UCLA applied-micro seminars, UCLA International Trade Seminars, the UCSB Applied Microeconomics Lunch, the USC-UCI-UCLA Urban Research Symposium, the Warwick Ph.D. Economics Conference, the LACAE, and the All California Labor Economics Conference at UCSC for helpful comments. This project was supported in part by the UCLA Ziman Center's Rosalinde and Arthur Gilbert Program in Real Estate, Finance and Urban Economics and the California Center for Population Research at UCLA with the grant (T32HD007545; P2CHD041055) from the NICHD. The content is solely my responsibility and does not represent the official views of the NICHD and the NIH. All errors are mine. This paper was previously circulated with title "The Valuation of Local Government Spending: Gravity Approach and Aggregate Implications".

†Department of Economics, Southern Methodist University. Address: 3300 Dyer Street, Suite 301, Dallas, TX 75205. Email: wookunkim@smu.edu.

1 Introduction

Public spending is unevenly distributed across cities. This spatial variation results due to the differences in the tax base and public finance system, which often redistributes local tax revenues across space. The uneven distribution of local public spending influences the location decisions of people both directly by provisioning local public goods and improving local amenities and indirectly through its interaction with local economic conditions such as home prices. As a result, the spatial variation in public spending may exacerbate or alleviate the uneven distribution of economic activity across space. Thus, quantifying the welfare implications of a public finance system is important: How does a spatial redistribution of tax revenue impact overall welfare?; What is the optimal level of fiscal redistribution? There is limited quantitative evidence on the welfare consequences of spatial fiscal redistribution. Furthermore, answering these questions requires an understanding of how much people value local government spending, on which the previous literature has little evidence.

In this paper, I estimate how much people value local government spending and examine the welfare consequences of spatial fiscal redistribution. To guide my analysis, I develop a quantitative spatial general equilibrium model, which embeds a simple public finance system with local taxation and spatial redistribution. A novel feature of the model is to allow spatial mobility of workers along two margins: commuting and migration. While commuting across space implies that workers need not work and live in the same location but bear the cost of commuting, migration together with commuting creates additional spatial frictions associated with migration and job finding. An important implication of the model is the gravity equation, which simultaneously captures both migration and commuting decisions. Combining the gravity equation with data from South Korea and a quasi-natural experiment of tax changes, I implement a new estimation approach to identify key elasticities of the model and shed light on the valuation of local government spending and the extent of rivalry associated with public goods in the spirit of Tiebout (1956). Having estimated the model, I evaluate the welfare consequences of redistributing public spending across space and numerically compute the optimal extent of fiscal redistribution (i.e., the share of local tax revenue appropriated for spatial redistribution). Throughout the paper, I show that accounting for both migration and commuting is crucial when estimating key elasticities that govern the spatial distribution of economic activity and, thus, the general equilibrium implications of counterfactual policy experiments.

In the model, workers are heterogeneous in terms of initial residence and choose where to live and where to work based on wages, cost of living, amenities, local public goods, and spatial frictions. There are three types of spatial frictions: the costs of migration, commuting, and job finding. Accounting jointly for both margins of mobility is important because migration and commuting decisions are linked via residential location decisions. Workers may move to places with generous provision of local government goods, but they may also find places attractive to live in if these places facilitate better access to jobs via commuting. Furthermore, workers from different origins may find some residential locations more or less costly to migrate to and some workplace locations difficult to find jobs in due to spatial frictions. Firms decide where to locate and how much to produce by choosing labor and floor space as inputs based on local productivity and factor prices. The supply of floor space is endogenously determined for commercial or residential use. Lastly, there exists local public spending, which is determined based on the local income taxes and spatial fiscal redistribution.

The model features agglomeration and dispersion forces. On the one hand, local public spending serves as an agglomeration force as higher residential density conditional on income implies a greater level of local public spending. On the other hand, concentration of residential density increases home prices, and the extent of the benefits people enjoy from the same level of local government spending may differ based on residential density due to the rivalry associated with the consumption of local public goods. In addition to these intra-regional spillovers, the spatial redistribution of tax revenue generates inter-regional fiscal externalities. The gravity equation derived from the model summarizes how the spatial mobility of workers are influenced by these forces. I apply data from South Korea and quasi-natural policy reforms to estimating the gravity equation.

The empirical setting of this paper is South Korea, where three key aspects 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 that provisions local public goods to its residents. This local spending is financed via income tax from the residents, part of which is locally retained while the rest is redistributed across districts. Second, the national tax policy reforms in 2008 and 2012 changed income tax rates, which I leverage as a source of exogenous variation in local government spending, residential density, and home prices to estimate the key reduced-form elasticities governing workers' spatial mobility. Third, a key data requirement to apply the model is to observe the joint distribution of migration and commuting patterns. Population Census of South Korea is a rare data source, which allows me to simultaneously observe both bilateral migration and commuting decisions. Based on the 2005, 2010, and 2015 Population Census, I construct a geo-coded panel data set of the number of household heads in terms of three locations: residence from 5 years ago, current residence, and workplace location.

I jointly estimate the reduced-form elasticities of worker's spatial mobility with respect to local public spending, residential density, and home prices based on the gravity equation. A rich set of fixed effects controls for a number of confounding factors including labor market returns and spatial frictions. Nonetheless, there remain endogeneity concerns due to unobservable time-varying residence-specific factors that are correlated with local public spending, residential density, and home prices. I propose an instrumental variable approach based on the national tax policy reforms, together with historical residential density to generate exogenous variation. The changes in the national income tax rates shift local government spending, home prices, and residential density, as workers are pulled in by local government spending and pushed out by increased home prices, while

[1] Districts in this paper are the 222 smallest administrative units in South Korea called *Si*, *Gun*, or *Gu*. They are smaller than the average U.S. counties in terms of land area. To give a sense of the scale, the total land area of South Korea is about 1% of the U.S. or about the same size as the state of Kentucky, U.S.

conditioning on income and the costs of spatial mobility with the fixed effects.

I find that workers are 1.07% more likely to choose to live in a district if its local government spending increases by 1%. A 1% increase in residential density and home prices decrease the probability of workers choosing this residence by 0.8% and 0.5%, respectively. Based on the structural relationship between the reduced-form elasticities, I estimate the extent of the rivalry associated with the consumption of local public goods. I find that the benefits of tax contribution by one resident to the local public spending is shared with the other residents; however, by holding local public spending fixed, workers benefit less from local public goods as residential density increases.

I estimate the effects of spatial frictions and wages on worker's spatial mobility. In line with the literature (e.g., Bryan and Morten 2019; Monte et al. 2018), the estimated distance elasticities of migration, commuting, and job finding are negative, statistically significantly different from zero, and stable over the sample period. I show that there are large, economically meaningful biases to the estimated distance elasticities of migration and commuting when only one margin is considered. For instance, the estimated distance elasticity of migration without considering commuting is biased upward toward zero because workers are willing to migrate over a longer distance when their destination offers better labor market access.[2] Estimating the distance elasticity of commuting while not accounting for migration leads to overestimation. The bias in this case may arise because a commuting decision often incurs costs of migration and job finding in addition to the direct cost of commuting between residence and workplace location.[3]

Following the approach in Ahlfeldt et al. (2015), I estimate how much people move in response to wages (i.e., the Fréchet shape parameter). Combining this estimate and the estimated reduced-form elasticities, I find that workers on average value an additional dollar of local government goods equal to 75 cents of their after-tax income when evaluated based on the 2015 average values of local government spending and income. Taking both margins of migration and commuting is crucial to estimate people's valuation of local government spending. This is because the instrumental variables constructed based on the income tax policy reforms would not be valid when only one of these two margins of mobility is considered.

Using the estimated model, I conduct a set of counterfactual policy experiments to shed light on the optimal organization of the local public finance system. The main objective of the experiments is to investigate the welfare consequences of spatial fiscal redistribution. Many countries around the world (e.g., Canada, Germany, Australia and Japan) make fiscal transfers across regions, which is similar to the South Korean system featured in this paper. I allow for counterfactual regimes to mimic what is observed in other countries with differing levels of redistribution (i.e., how much

[2] For example, Bryan and Morten (2019) estimate the distance elasticity of migration based on the migration patterns in the U.S. and Indonesia without taking commuting into account. Similar to their approach, I estimate the distance elasticity of migration without accounting for commuting patterns, and the estimate I obtain is close to their estimate. However, once both migration and commuting patterns are taken into account, the magnitude of the same elasticity increases by almost 5 times.

[3] Ahlfeldt et al. (2015) estimate the distance elasticity of commuting in the context of the city of Berlin, Germany. When following their estimation strategy and using commuting patterns in isolation (without allowing for a migration margin), I find a commuting elasticity similar to theirs. However, this estimate is about 56% larger in magnitude (more negative) than the estimate based on both migration and commuting.

local government spending depends on redistributive intergovernmental transfers relative to local taxation). A lower extent of redistribution implies local governments retain a higher share of their local income tax revenue and contribute less to fiscal redistribution. In the extreme case, local government revenue is equal to local income tax revenue with no intergovernmental transfers. The optimal level of redistribution depends on the interplay of intra- and inter-regional fiscal spillovers and the spatial distribution of economic activity.

I find that overall welfare increases at most by 0.12% if the national government collected a smaller fraction of local tax revenue for redistribution, compared to what is observed in 2015. Completely eliminating the redistributive intergovernmental transfers, however, leads to a large aggregate welfare loss of 1.2%. The results indicate that transfers of income from fiscally strong districts (i.e., districts with higher average income) to the weak are too much under the observed redistribution policy in 2015. The benefits of the transfers in the net-receiving districts are dominated by the losses in the net-contributing districts. In addition, I show that different assumptions on spatial mobility of workers made in the literature call for a significantly different extent of redistribution. To demonstrate, 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, the cost of commuting is the only source of spatial friction. As a result, workers can afford to commute longer from net-receiving districts where they would enjoy a greater level of local government spending than how much they contribute. To mitigate this tendency, the optimal redistribution erroneously calls for a significantly lower extent of redistribution.

My paper builds upon several existing literature. This paper contributes to a growing literature on quantitative economic geography models.[4] There are a number of recent papers that have studied the migration and commuting decision, but these have done so only separately. In the case of migration, Bryan and Morten (2019) 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.[5] One of the common assumptions in this literature is that people live and work in the same locations. 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, Colombia, respectively. The literature on commuting assumes often implicitly zero spatial frictions associated with migration and job finding.[6] To the best of my knowledge, this paper is the first to present a spatial equilibrium model that features both bilateral migration and commuting in the economic geography literature. By doing so, I show that considering both margins of mobility at the same time is crucial when estimating key elasticities governing the spatial mobility of workers. Furthermore, few papers in

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

[5]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).

[6]Monte et al. (2018) have a notion of migration in addition to commuting; however, they assume idiosyncratic costs of migration as opposed to the systematic costs of migration based on distance in this paper.

this literature study the roles of the public sector, which serves as a source of endogenous local amenities.

The urban and public economics 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). In the spirit of Rosen (1979) and Roback (1982), some papers estimate positive amenity values for government spending and regulations from housing prices (Cellini et al., 2010; Black, 1999; Chay and Greenstone, 2005).[7] In addition, there are only a few papers that directly estimate how much workers value government spending. Using a spatial general equilibrium framework, Suárez-Serrato and Wingender (2014) estimate the effect of federal spending on local economies in the U.S. by exploiting changes in population levels that are 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.

My approach includes various novel features relative to the previously mentioned 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 in the U.S.). Given the granular spatial units, I leverage both migration and commuting patterns to estimate how much workers value local government spending. 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) and evaluate the extent of rivalry associated with local public goods. Fourth, I evaluate the welfare consequences of spatial redistribution of public spending and shed light on the optimal redistribution policies.

Lastly, this paper also contributes to the literature on fiscal decentralization. The majority of the papers in this literature focus on the theoretical and empirical examination of the consequences of the changes in fiscal autonomy of local governing entities.[8] 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 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

[7] 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.

[8] 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.

equalization policy in Canada. I contribute to the literature on fiscal decentralization by computing the optimal balance between local taxation and redistribution while taking the spatial mobility of workers into account. Furthermore, I show that assumptions on spatial mobility of workers give rise to different optimal policies.

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 the 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 estimates reduced-form elasticities measuring the distance-elasticities of migration, commuting, and job finding. 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. I quantify the welfare consequences of spatial fiscal redistribution 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 to quantify the aggregate welfare consequence of fiscal redistribution across space. Specifically, 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 the National Tax Reforms in 2008 and 2012, which serve as a source of exogenous variations to estimate key elasticities of spatial mobility.

2.1 Data

The observed spatial distribution of workers is a consequence of decisions on two margins of the 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, local government spending should vary at the same spatial disaggregation across which workers actively make both migration and commuting decisions. South Korea is one of a few countries that meet all the requirements. Third, to evaluate aggregate implications, the data has to be spatially representative.

The data source for the spatial distribution of workers is the restricted-access 2005, 2010, and 2015 Population Census of South Korea.[9] 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.[10] The sample

[9]The Population Census of South Korea is conducted every five years and sample 20% of the entire population.

[10]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% of the households in the Population Census have male as household heads. The female labor force participation in South Korea is one of the lowest among the OECD countries (Lee, 2017). The

size is about 3.5 million households. The Census questionnaire asks the district of residency five years ago, the current district of residence, and the district of workplace location. Based on this information, I construct a panel data set of three locations (i.e., districts of residence five years ago, current residence, and workplace location) that captures the spatial distribution of workers jointly in terms of migration and commuting.

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 collected the total revenue and revenue from different sources: local income taxes and intergovernmental transfers. This information allows me to recover the share of the intergovernmental transfer that each locality received from the national government in a given year. In addition, the Ministry of Land, Infrastructure, and Transport publishes the land price index at the district level. I collected this information for 2005, 2010, and 2015. The index is defined as the average land price in a given year, normalized against the average land price in 2004.

I supplement the main data set with local characteristics in 2015 using the administrative 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.[11] 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 wage 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 the distance between every centroid of the districts.[12]

2.2 Spatial Mobility in South Korea

While South Korea is only about 1% of the U.S. geographically, the population level was about 51 million in 2015, which was about 16% of the population in the U.S. The spatial unit used in this paper is districts in South Korea that correspond to the smallest administrative units with local governing authority. I will hereon refer to these district-level governing entities as local governments, and I focus on the 222 contiguous districts that partition the South Korean mainland, excluding the districts of Jeju Island.[13] The average size of each district is 224,310 in terms of population (91,471 households), approximately twice as large as the average population of a county in the U.S.

I describe the commuting and migration patterns in South Korea. First, workers in South Korea spend about 7.3% of their workday commuting between their residence and workplace locations,

age restriction is to only include workers who have completed education and the compulsory military duties. Also, less than 1% of workers report a one-way commuting distance over 90 kilometers.

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

[12] 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.

[13] I also exclude a few districts that split or merged during the sample period.

Table 1: Summary Statistics

	(1) Observation	(2) Mean	(3) Std. Dev.	(4) Min	(5) Max
A. Commuting Patterns					
Commuters from Residence	666	0.273	0.237	0.000	0.773
Commuters to Workplace	666	0.296	0.212	0.000	0.916
B. Migration Patterns					
In-Migrants	666	0.187	0.079	0.053	0.559
Out-Migrants	666	0.180	0.071	0.048	0.443
C. Local Government Budget					
Per-Capita Local Gov't Expenditure	666	3.309	2.763	0.316	15.000
Total Local Gov't Expenditure	666	362,785	233,524	59,614	1,881,082
Intergovernental Transfers	666	242,517	128,348	24,341	799,009
Local Income Tax Revenue	666	64,067	95,793	1,663	779,143

Notes: This table reports the summary statistics for 222 districts in 2005, 2010, and 2015. The data set 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. The variable "Commuters to Workplace" measures the fraction of workers employed in a district who commute from other districts. Similarly, the variables "In-Migrants" and "Out-Migrants" measure the fraction of residents of a district who moved in within 5 years from another district 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).

similar to the commuting patterns documented in Monte et al. (2018) for the U.S. and Schafer (2000) for 26 countries around the world.[14]

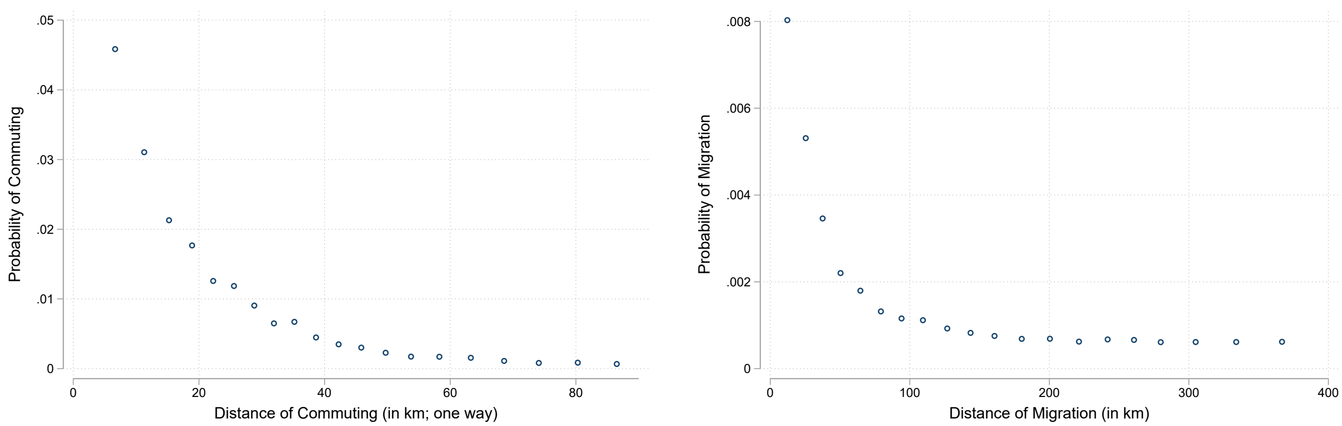
In Panel A of Table 1, I report summary statistics on the commuting patterns in 2005, 2010, and 2015. On average, about 27% of residents work outside their district of residence and about 30% of workers commute to work from other districts. In addition, I plot the fraction of residents commuting to other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

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

Figure 1: Commuting and Migration vs. Distance

(a) Commuting (b) Migration



Note: This figure shows that the probabilities of commuting (left panel) and migration (right panel) decrease as distances of commuting and migration increase. The probabilities are computed using the Population Census of South Korea (2005, 2010, and 2015). Each point corresponds to the mean probability for each 5th percentile of commuting and migration distances.

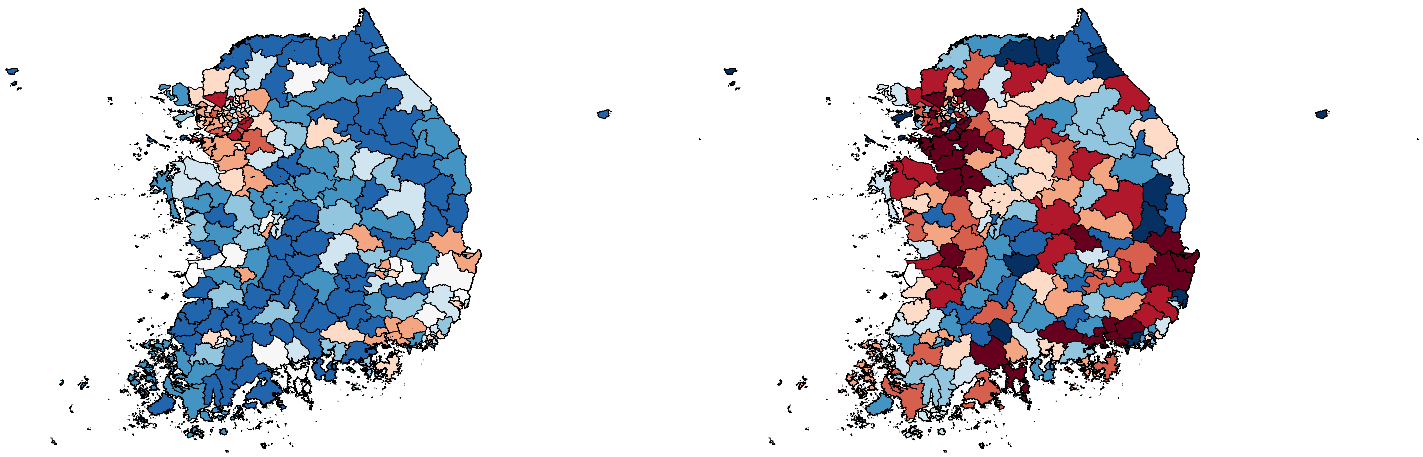
annually; the implied annual inter-district migration rate is around 13%.[15] Aggregated at the province level (i.e., 16 groups of districts), the annual migration rate in South Korea is 5%, similar to the inter-county migration rate in the U.S. (Molloy et al., 2011). In Panel B of Table 1, I report the probabilities of migration for over 5 years at the district level. On average, about 19% of residents in a district are migrants who have migrated from other districts within 5 years, while 18% 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, the probability of migration decreases with distance (Bryan and Morten, 2019; Morten and Oliveira, 2018).

In Figure 2, I plot the number of households in the figure on the left and the total local expenditure on the right by district. There are districts with generous local expenditure and many households. This pattern may suggest that workers are more likely to reside in districts that are associated with generous provision of local government goods and services or that the local government spending is higher in places with more workers simply because of larger tax bases. It is important to note that the key residential determinants of migration and commuting such as wages, home prices, and amenities are interlinked with each other. For instance, greater local government spending would increase local population by improving local amenities, while the increase in residential density would push up home prices as the demand for housing rises and local

[15] 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.

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

(a) Residential Density (b) Local Government Spending



Note: This figure plots the spatial distribution of workers in terms of their residences (left panel) and local government spending (right panel) observed in 2015. Red (blue) districts indicate higher (lower) values.

government spending as its tax base increased. 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 Public Finance System and National Tax Reforms

The total local government expenditure accounts for about 8% 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% and 72% of local government spending, respectively.[16] 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 363 million USD; the average per-capita local expenditure is 7,638 USD, widely ranging from 906 USD to 29,622 USD.[17] There is substantial spatial variation in the degree to which districts depend on the local income taxes and intergovernmental transfers for their spending. [18] The share of local government revenue from local income tax ranges from 2.1% to 56%.

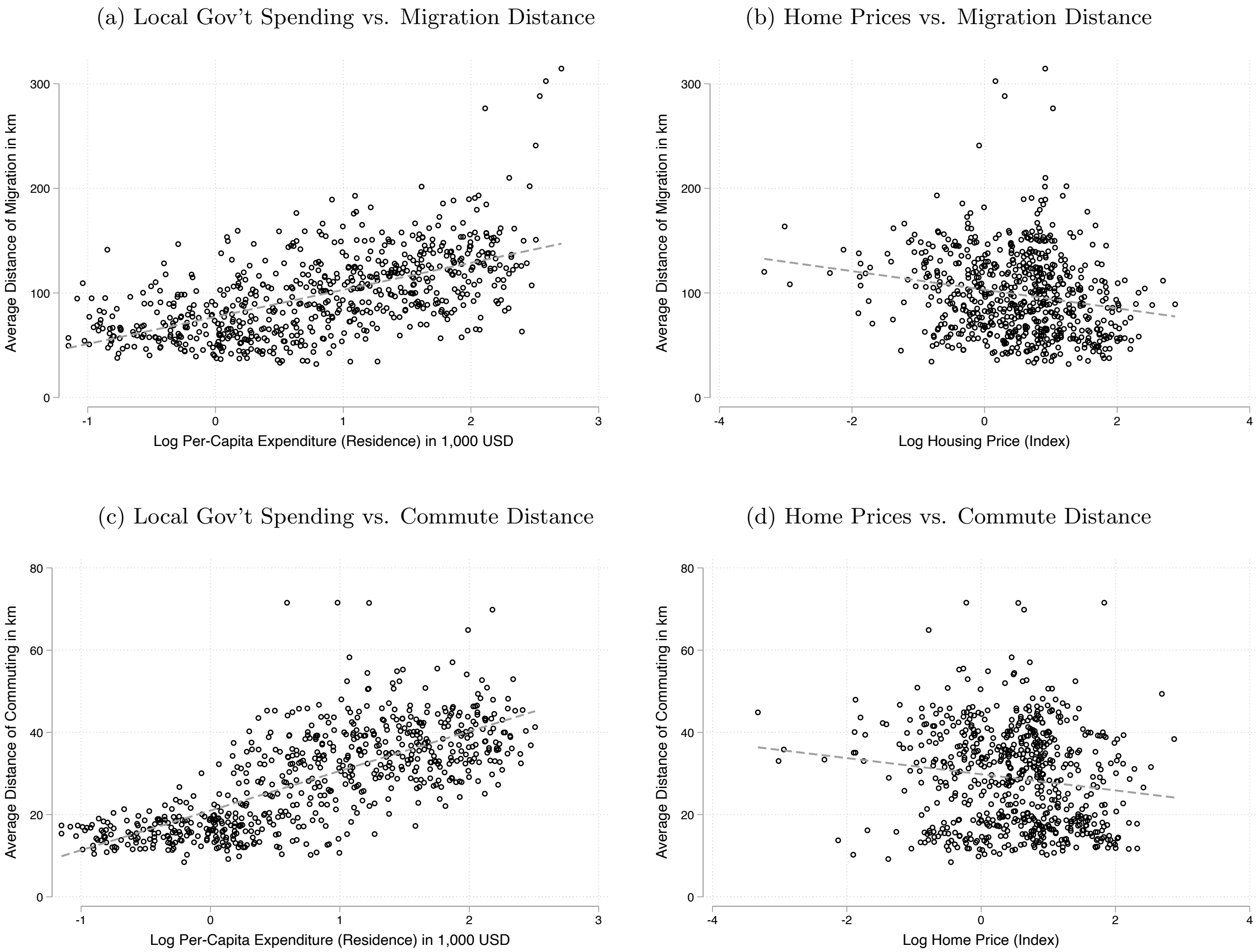
The national fiscal policies consist of a progressive income tax system and extents of fiscal decentralization and redistribution. Together with the national policies, local tax bases (number of workers and their income) determine local government revenue. The Local Autonomy Act—first enacted in 1949—was revived in 1991 after 30 years of suspension due to military dictatorships

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

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

[18]Figure B.1 in Appendix shows the spatial distribution of local government revenue.

Figure 3: Distances of Migration and Commuting vs. Local Government Spending and Home Prices



Note: This figure shows the raw correlation between how far workers migrate and commute and local government spending/home prices. Each observation is a district-year pair. The panels 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 panels 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).

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 as well as the Income Tax Act promulgate 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 national income tax rates that are outlined in the Income Tax Act. The income tax rates are progressive and uniformly applied 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 *the extent of redistribution* and the amount of income tax revenue left at the local level as *local income tax*. Lastly, the national government rebates *intergovernmental transfers* back to local governments. The amount of intergovernmental transfers that each district receives is equal to a certain share of the total tax revenue delivered to the national government. I refer to the shares specific to each district as *the rules of 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.[19] In general, the rules of redistribution favor low-amenity areas where extra public funding would improve the standard of living.

Tax Reforms

There were two major reforms on the national tax rates: one in 2008 and the other in 2012, both of which took place between the sample periods focused on in this paper. In 2008, the Income Tax Act was amended to decrease income tax rates across income brackets: from 11% to 8.8% for the low income bracket (annual income less than 12,000 USD); from 22% to 18.7% for the middle income group (12,000 to 46,000 USD); and from 33% to 28.2% for the high income group (46,000 to 88,000 USD).[20] 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-income group, and to 26.4 percent for the high-income group. The tax reforms, however, did not affect the extent and rules of redistribution outlined in the Local Subsidy Act.[21]

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, and local government goods and services. My model is different

[19] The Local Subsidy Act details the formula employed to determine the rules of redistribution 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). The rules of redistribution favor districts with lower amenity values and higher population density. 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).

[20] 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% 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 previous change was made back in 1994. Since the first reform, the cutoffs for the lowest three income brackets have remained the same.

[21] Figure B.2 in the Appendix compares the rules of distribution in 2005, 2010, and 2015. The estimated slopes comparing the rules of distribution in a given year to these five years ago are close to 1.

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 simultaneously features both commuting (Ahlfeldt et al., 2015; Tsivanidis, 2019) and migration decisions (Bryan and Morten, 2019; 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.

3.1 Model Environment

Workers are born in or assigned to initial residence or origin o. The initial distribution of workers is given by π_o. I assume each worker inelastically supplies one unit of labor. The whole economy has R measure of workers (interchangeably referred to as residents) and comprises of J arbitrary, discrete number of spatial units, indexed by current residence r , workplace location m , and origin o. 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. 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.[22] 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. After observing an idiosyncratic utility shock for every possible pair of residence r and workplace m, each worker chooses a residence-workplace pair that maximizes her utility given her origin (initial residence), after-tax wages, floor space prices, local government goods and services, local amenities, and the iceberg costs of mobility.

3.2 Worker's Location Decisions

The preferences of 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 worker i who chooses to move from origin 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:

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

[22]Given the national income tax rates $\tau(\cdot)$, the wage from workplace location m determines income tax rate $\tau(w_m) \equiv \tau_m$.

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 Albouy (2012) and Fajgelbaum et al. (2019), real government expenditure enjoyed by each worker living in district r g_r is equal to 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 the 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 worker i from origin o choosing to live in district r and commute to district m is $V_{iorm} = z_{irm} v_{orm}$,where

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

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 $D_{orm} = \varepsilon_{orm} D_{or} D_{rm} D_{om}$. Following the standard approach in the spatial literature, I further define the composite iceberg cost rising from migration, commuting, and job finding D_{orm} as follows:

$$D_{orm} = \varepsilon_{orm} \exp(\rho d_{or} + \kappa d_{rm} + \delta d_{om}), \tag{4}$$

where d_{jk} is the distance between district j and k; the parameter ρ controls the size of migration cost with respect to distance between previous and current residences d_{or}; the parameter κ controls the size of commuting cost with respect to distance between residence and workplace d_{rm}; the parameter δ controls the size of job finding cost with respect to distance between previous residence and workplace d_{om}; ε_{orm} is an independent stochastic error following a log normal distribution with mean equal to 1. The composite iceberg cost enters the indirect utility function multiplicatively. Therefore, there is an isomorphic formulation in which after-tax wages and amenities values are reduced due to the costs of migration, commuting, and job finding costs.

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

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

model this heterogeneity in preference in spirit of McFadden (1974) and Eaton and Kortum (2002). This preference shock is drawn from an independent Fréchet distribution:

$$\Pr(z_{irm} < z) = \exp(-T_r M_m z^{\epsilon}), \tag{5}$$

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 $\epsilon > 1$ 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 is given by:

$$\pi_{rm|o} = \frac{\left(\frac{\tilde{B}_r(1-\tau_m)\tilde{w}_m}{\varepsilon_{orm}\exp(\rho d_{or}+\kappa d_{rm}+\delta d_{om})Q_r^{1-\beta}}\left(\frac{G_r}{R_r^{\theta}}\right)^{\lambda}\right)^{\epsilon}}{\sum_{r'=1}^{J}\sum_{m'=1}^{J}\left(\frac{\tilde{B}_{r'}(1-\tau_{m'})\tilde{w}_{m'}}{\varepsilon_{orm}\exp(\rho d_{or'}+\kappa d_{rm'}+\delta d_{om'})Q_{r'}^{1-\beta}}\left(\frac{G_{r'}}{R_{r'}^{\theta}}\right)^{\lambda}\right)^{\epsilon}} \equiv \frac{\Phi_{orm}}{\Phi_o}, \text{where } \Phi_o = \sum_{r'=1}^{J}\sum_{m'=1}^{J}\Phi_{orm}. \tag{6}$$

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 residential locations with high amenity values and local government expenditure and lower per-unit floor space price and the extent of rivalry.[25] Workers are more likely to commute to workplace locations with higher after-tax wages. Lastly, the interplay of the costs of migration, commuting and job finding governs the probability of workers living in r and working in m conditional on their origin.

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 previous section. 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 as a source of exogenous variation in local government spending, home prices, and residential density. In Section 4.2 and 4.3, I present the estimation results and discuss the

[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 the gravity equation (6) is derived. 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, 2019; Morten and Oliveira, 2018; Monte et al., 2018; Moretti and Wilson, 2017). In Appendix D.2, I show the theoretical correspondence between the gravity equation presented in this paper and the gravity equations used in the migration and commuting literature.

interpretation and robustness of the estimated reduced-form elasticities.

4.1 Estimation Strategy

The gravity equation (Eq. 6) describes how workers sort across districts in terms of residential and workplace locations from their previous residences. 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:

$$\ln \pi_{orm,t} = \phi_{om,t} + \phi_{or} + \phi_{rm} + \underbrace{\lambda\epsilon}_{\beta_G} \ln G_{r,t} - \underbrace{\theta\lambda\epsilon}_{\beta_R} \ln R_{r,t} - \underbrace{(1-\beta)\epsilon}_{\beta_Q} \ln Q_{r,t} + \zeta_{orm,t}. \qquad (7)$$

The coefficients in front of log local government spending ($\beta_G = \lambda\epsilon$), log residential density ($\beta_R = \theta\lambda\epsilon$), and log prices of floor space ($\beta_Q = (1-\beta)\epsilon$) are the reduced-form elasticities 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., the initial distribution of workers (6)) 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 and workplace location characteristics that make it a more or less attractive place to live and/or work in.[26] Lastly, the error term $\zeta_{orm,t}$ includes the rest of the factors in equation 6 (i.e., time-varying adjusted amenities and stochastic components of the iceberg costs).

The errors in Equation 7 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). 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.

Fixed Effects

The mapping between the econometric specification (Eq. 7) and the gravity equation from the spatial model (Eq. 6) helps to understand potential confounders and consequent biases. First, the job finding fixed effects interacted with year dummy variables $\phi_{om,t} = \ln(1-\tau_{m,t})^{\epsilon}\tilde{w}^{\epsilon}_{m,t}\exp(-\delta\epsilon d_{om})\pi_{o,t}/\Phi_{o,t}$ 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

[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.

from a workplace location, workers may be 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 that differs by origin, the OLS estimate of β_G will be downward biased.

Next, a higher net labor market access attracts residents. This positive correlation between the residential density and the labor market return biases the OLS estimate of $-\beta_R$ 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 push 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 $-\beta_R$ and $-\beta_Q$ would be biased downward.[27]

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 time-varying values of adjusted residential amenities. Local government spending and local amenities are, in this empirical setting, 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 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 $-\beta_R$ would be based upward. Lastly, high amenity values would be priced into home prices in the sense of hedonic pricing. Then, the OLS estimate of $-\beta_Q$ would suffer from an upward bias toward zero.

Furthermore, there is an additional concern of measurement error with respect to Q_r. I do

[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.

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 the endogenous regressors are correlated with each other, all the other OLS estimates would be also biased.

Instrumental Variables

Because the estimating equation (7) has three endogenous variables, 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:

$$IV^{b}_{r,t} = \tau_{b,t}\pi_{b|r,2000}, \tag{8}$$

where $\tau_{b,t}$ denotes the national income tax rates in year t for income group b (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 b. Because I do not observe the income distribution at the district level, I use the education attainment level to proxy for income. I construct $\pi_{b|r,2000}$: the share of workers living in district r in 2000 by low (i.e., high school) and high (i.e., some college degrees) educational levels.

The instrumental variables ($IV^{low}_{r,t}$ and $IV^{high}_{r,t}$) capture the tax contributions of low and high income groups predicted by predetermined distribution of education attainment level of residents 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, home prices, and residential densities. There are two sources of variation in the proposed instruments. One 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 were not correlated with the changes in the educational composition in each district. In Appendix B.4, I provide evidence 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 robust to potential sorting by educational levels.

Second, the last instrumental variable $IV^R_{r,t}$ is based on the historical residential density as previously used in Ciccone and Hall (1996) and de la Roca and Puga (2017). I use the log 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 characeteristics that affect worker mobility because the changes in the type of economic activity and historical events like war reshape the economic landscape.[28] The validity of the instrument hinges on the assumption that historical residential densities do not directly affect the worker location decisions today.[29] 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:

$$E\begin{bmatrix} IV^{low}_{r,t}\zeta_{orm,t} & |\phi_{om,t},\, \phi_{or},\, \phi_{rm} \\ IV^{high}_{r,t}\zeta_{orm,t} & |\phi_{om,t},\, \phi_{or},\, \phi_{rm} \\ IV^{R}_{r,t}\zeta_{orm,t} & |\phi_{om,t},\, \phi_{or},\, \phi_{rm} \end{bmatrix} = 0. \tag{9}$$

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 changes its sign and increases to 0.097. This increase can be explained by netting out the negative correlation between local government expenditures and labor market returns as local government spending in regions with lower labor market access is higher due to 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 residential density and after-tax wages discounted by the cost of job finding in line with intuition. Lastly, the OLS estimate of β_Q increases to -0.01 closer to zero; however, this estimate is statistically indistinguishable from zero. On the one hand, the increase in the estimate of β_Q is against the direction of bias associated with omitting

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

[29] 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).

Table 2: OLS Estimates of Elasticities of Worker Mobility

Dependent Variable: $\ln \pi_{orm,t}$	(1)	(2)	(3)	(4)
Local Gov't Expenditure, $\ln G_{r,t}$ ($\beta_G = \lambda\epsilon$)	-0.231^{***}	0.096^{**}	0.096	0.087^{***}
	(0.015)	(0.037)	(0.127)	(0.027)
Number of Households, $\ln R_{r,t}$ ($\beta_R = \theta\lambda\epsilon$)	0.120^{***}	0.063^{**}	0.298	0.593^{***}
	(0.012)	(0.025)	(0.196)	(0.047)
Floor Space Price, $\ln Q_{r,t}$ ($\beta_Q = (1-\beta)\epsilon$)	-0.042^{***}	-0.011	0.006	-0.002
	(0.013)	(0.031)	(0.030)	(0.006)
Observations	257,174	257,174	257,174	257,174
Fixed Effects:				
Job Finding Pair × 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 mobility to local government expenditure and resident population levels based on Eq. 7, starting with a simple estimate without any fixed effects in Column (1) and gradually adding the fixed effects discussed in Section 4.1. Column (4) corresponds to Eq.Eq. 7 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×year, commuting pair×year, and a triplet of previous and current residences and workplace: $***p < 0.01$, $**p < 0.5$, $*p < 0.1$.

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.[30]

[30] The estimates in Column (4) is based on the fully saturated specification (7). 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

Table 3: 2SLS Estimates of Elasticities of Worker Mobility

	(1)	(2)	(3)	(4)	(5)
	OLS	First Stage	First Stage	First Stage	2SLS
Dependent Variable:	$\ln \pi_{orm,t}$	$\ln G_{r,t}$	$\ln R_{r,t}$	$\ln Q_{r,t}$	$\ln \pi_{orm,t}$
Local Gov't Expenditure, $\ln G_{r,t}$ ($\beta_G = \lambda\epsilon$)	0.087***				1.033***
	(0.027)				(0.356)
Number of Households, $\ln R_{r,t}$ ($\beta_R = \theta\lambda\epsilon$)	0.593***				−0.807
	(0.047)				(0.577)
Floor Space Price, $\ln Q_{r,t}$ ($\beta_Q = (1-\beta)\epsilon$)	−0.002				−0.489***
	(0.006)				(0.063)
Predicted Tax Contribution (low), $IV^{low}_{r,t}$		13.893***	5.623***	60.322***	
		(1.344)	(1.016)	(6.101)	
Predicted Tax Contribution (high), $IV^{high}_{r,t}$		14.118***	6.762***	21.080***	
		(0.742)	(0.608)	(2.902)	
Number of Households 30 years ago, $IV^{R}_{r,t}$		0.028***	−0.014**	0.108***	
		(0.008)	(0.006)	(0.027)	
Observations	257,174	257,174	257,174	257,174	257,174
SW F-stat		46.737	42.677	57.073	
$\hat{\theta}$	−6.802				0.780***
	(5.381)				(0.093)

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 Eq. 7. Column (1) is identical to Column (4) in Table reftable:OLS-G-R. 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 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: $***p < 0.01$, $**p < 0.5$, $*p < 0.1$. I estimate the congestion parameter θ based on the structural relationship between the estimated reduced form parameters ($\beta_R/\beta_G = \theta$): $\hat{\theta} = 0.787\,(0.199)$.

Table 3 summarizes the two-stage least squares estimation results in Column (2) to (5) and compares the results with the OLS results in Column (1). First, according to the estimates 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 in Column (4). 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 1.03 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 0.807 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 0.78 with a standard error equal to 0.04. Based on these estimates, I can reject the null hypothesis that θ lies outside of its theoretical range between 0 and 1. 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 -

local amenity values that make residence more attractive and are correlated with the included regressors.

0.489, 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 0.489 percent as home prices increase by 1 percent.

4.3 Interpretation of the Estimates

The elasticity of worker's mobility to government expenditure has not been extensively estimated in the previous literature. One exception is Suárez-Serrato and Wingender (2014); they 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 1.03 is close to their estimates.[31]

The literature on agglomeration economies includes population density a determinant of endogenous amenities and productivity (Ciccone and Hall, 1996; Glaeser and Maré, 2001; Ahlfeldt et al., 2015; de la Roca and Puga, 2017; Almagro and Domínguez-Iino, Almagro and Domínguez-Iino). 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.[32] 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.

In Table 4, 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 is violated in each case. Based on the migration patterns 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. Based on the commuting patterns alone, the same effect is substantially overestimated because, in addition to direct cost of commuting, there exist migration and job finding costs that enable each commute. This set of results emphasizes the importance of jointly accounting for both migration and commuting and for proper conditioning to

[31] 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 effect of government expenditures that affect firms as well as residents. In my setting, I am able to focus on the effect of local government spending on the residents.

[32] 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 β_R in Eq. 7 is $(\theta\lambda - \gamma)\epsilon$ instead of $\theta\lambda\epsilon$.

Table 4: Estimation Results based on Migration and Commuting Flows

	(1)	(2)	(3)	(4)	(5)	(6)
	Both		Migration Flows		Commuting Flows	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	$\ln \pi_{orm,t}$	$\ln \pi_{orm,t}$	$\ln \pi_{or,t}$	$\ln \pi_{or,t}$	$\ln \pi_{rm,t}$	$\ln \pi_{rm,t}$
Local Gov't Expenditure, $\ln G_{r,t}$ ($\beta_G = \lambda\epsilon$)	0.087***	1.033***	0.353***	−1.341**	0.288***	2.139
	(0.027)	(0.356)	(0.042)	(0.576)	(0.050)	(1.706)
Number of Households, $\ln R_{r,t}$ ($\beta_R = \theta\lambda\epsilon$)	0.593***	−0.807	1.120***	2.526***	1.118***	0.010
	(0.047)	(0.577)	(0.100)	(0.766)	(0.084)	(1.897)
Floor Space Price, $\ln Q_{r,t}$ ($\beta_Q = (1-\beta)\epsilon$)	−0.002	−0.489***	0.042***	0.407***	−0.031***	−1.094**
	(0.006)	(0.063)	(0.009)	(0.053)	(0.011)	(0.524)
Observations	204,576	204,576	58,995	58,995	19,121	19,121
Fixed Effects:						
$\phi_{om,t}$, ϕ_{or}, ϕ_{rm}	Y	Y	N	N	N	N
$\phi_{o,t}$, ϕ_{or}	N	N	Y	Y	N	N
$\phi_{m,t}$, ϕ_{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 home 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.

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. These frictions not only affect the spatial distribution of economic activity, but also interact with local determinants of spatial sorting. 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 estimating the distance-elasticities of migration and commuting.

5.1 Spatial Frictions in Migration and Commuting Decisions

I rewrite the gravity equation (Eq. 6) by grouping the location-specific factors by residence ϕ_r, by workplace location ϕ_m, and by previous residence ϕ_o:

$$\pi_{orm} = \frac{\phi_o \phi_r \phi_m}{\underbrace{(\varepsilon_{orm} D_{or} D_{rm} D_{om})}_{D_{orm}}{}^{\epsilon}}. \tag{10}$$

I refer to Eq. 10 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 Eq. 10, the expression for the spatial distribution of workers by their origins and current residences (migration flows) is given by:

$$\pi_{or} = \frac{\phi_o \phi_r}{D_{or}^{\epsilon}} \underbrace{\sum_{m=1}^{J} \frac{\phi_m}{(\varepsilon_{orm} D_{rm} D_{om})^{\epsilon}}}_{ALMA_{or}\varepsilon_{or}}. \tag{11}$$

The key difference between the expression above (11) 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 components 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. Therefore, $ALMA$ 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. [33]

Second, the literature on commuting employs a gravity equation, which summarizes the spatial distribution of workers in terms of their residential and workplace locations. By summing π_{orm} in Equation (10) over initial residences, I obtain a gravity equation that characterizes the commuting patterns of workers:

$$\pi_{rm} = \frac{\phi_r \phi_m}{D_{rm}^{\epsilon}} \underbrace{\sum_{o=1}^{J} \frac{\phi_o}{(\varepsilon_{orm} D_{or} D_{om})^{\epsilon}}}_{AMMA_{rm}\varepsilon_{rm}} . \tag{12}$$

Again, the key difference between the gravity equation above (Eq. 12) 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 when trying to do so from previous residence o. Similar to the direct cost of commuting, this indirect cost makes a commute less attractive.[34]

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:

$$D_{or} = \exp(\rho d_{or}),\ D_{rm} = \exp(\kappa d_{rm}),\ D_{om} = \exp(\delta d_{om}). \tag{13}$$

[33] 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.

[34] $AMMA$ measures how accessible each commute is for workers originating from different places on average and varies at the commute-pair level. 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.

The parameters ρ, κ, and δ control the sizes of migration, commuting, and job finding costs with respect to distance. 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 Eq. Equation (10) by adding time subscripts:

$$\pi_{orm,t} = \frac{\phi_{r,t}\phi_{m,t}\phi_{o,t}}{\varepsilon_{orm,t}^{\epsilon} \exp(\rho\epsilon d_{or} + \kappa\epsilon d_{rm} + \delta\epsilon d_{om})}. \tag{14}$$

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

Cost of Migration with respect to Distance

I take the log transformation of both sides of Eq. 14and obtain the expression as follows:

$$\ln \pi_{orm,t} = \phi_{rm,t} + \phi_{om,t} - \rho\epsilon d_{or} + \varepsilon_{orm,t}^{mig}, \tag{15}$$

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 $\varepsilon_{orm,t}^{mig}$ corresponds to the log of the stochastic error $\varepsilon_{orm,t}$; I assume this error term is orthogonal to distances of migration.[35] I allow the errors to be correlated across migration pairs.

Column (1) of Table 5 reports the estimation result. The estimated distance-elasticity of migration is -0.0328, which is statistically significant at the 1% significance level. This means one kilometer increase in the distance of migration decreases the probability of migration by 3.28 percentage points.[36] The magnitude of this estimate is much larger than the estimates in the literature. For example, Bryan and Morten (2019) estimate the elasticity of migration to distance in the U.S. (-0.553) and Indonesia (-0.717). Re-scaling my estimate of semi-elasticity by the average migration distance (75.34 kilometers), the implied elasticity of migration to distance based on my estimate is -2.47. The large difference between the estimates arise because the migration patterns alone do

[35] I estimate Equation (15) 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.

[36] In Appendix, Table B.4 presents the estimated distance-elasticity of migration based on Eq. 15 starting without any fixed effects and gradually adding them. In the last column, I show that the distance-elasticity does not vary across the sample periods.

not properly control for the fact that the values of local labor market access at each destination may differ by the origins of workers as shown in Eq. 11.

Table 5: Distance Elasticities of Migration, Commuting, and Job Finding

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	$\ln \pi_{orm,t}$	$\ln \pi_{or,t}$	$\ln \pi_{orm,t}$	$\ln \pi_{rm,t}$	$\ln \pi_{orm,t}$
Distance, d_{or} $(-\rho\epsilon)$	-0.0328^{***}	-0.0071^{***}			
	(0.0009)	(0.0001)			
Distance, d_{rm} $(-\kappa\epsilon)$			-0.0450^{***}	-0.0744^{***}	
			(0.0006)	(0.0009)	
Distance, d_{om} $(-\delta\epsilon)$					-0.0157^{***}
					(0.0003)
Observations	257,583	70,444	257,583	20,335	257,583
Fixed Effects:					
Commuting Pair × Year ($\phi_{rm,t}$)	Y	N	N	N	Y
Job Finding Pair × Year ($\phi_{om,t}$)	Y	N	Y	N	N
Migration Pair × Year ($\phi_{or,t}$)	N	N	Y	N	Y
Past/Current Residence × Year ($\phi_{o,t}$, $\phi_{r,t}$)	N	Y	N	Y	N
Residence and Workplace × Year ($\phi_{r,t}$, $\phi_{m,t}$)	N	Y	N	Y	N

Notes: In this table, I report the OLS estimates of elasticities of worker mobility to local government expenditure and resident population levels based on Eq. 7, starting with a simple estimate without any fixed effects in Column (1) and gradually adding the fixed effects discussed in Section 4.1. Column (4) corresponds to Eq.Eq. 7 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×year, commuting pair×year, and a triplet of previous and current residences and workplace: $***p < 0.01$, $**p < 0.5$, $*p < 0.1$.

To rule out the possibility that the large difference is an artifact of my data, I follow the standard strategy in the migration literature and estimate the semi-elasticity of migration based on the migration flows alone:

$$\ln \pi_{or,t} = \tilde{\phi}_{r,t} + \tilde{\phi}_{ot} - \rho\epsilon d_{or} + \varepsilon^{mig}_{or,t}, \tag{16}$$

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 $\varepsilon^{mig}_{or,t}$ 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 Eq. 11 with time subscripts on all the terms except distances, $\varepsilon^{mig}_{or,t}$ 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 Eq. 15 and report the estimated coefficient in front of distance in Column (2). Conforming to the expected direction of the omitted variable bias, the estimate is much smaller in magnitude as workers are willing to migrate longer distances when they face higher returns from the local labor market at the destination. Scaling this estimate by the average distance of migration, the implied distance-elasticity of migration in my setting is -0.535, close to the estimate of -0.553 in the case of U.S. (Bryan and Morten, 2019).

Cost of Commuting with respect to Distance

I estimate the semi-elasticity of commuting with respect to distance using the equation as follows:

$$\ln \pi_{orm,t} = \phi_{or,t} + \phi_{om,t} - \kappa\epsilon d_{rm} + \varepsilon^{com}_{orm,t}, \tag{17}$$

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 $\varepsilon^{com}_{orm,t}$, orthogonal to distances of commuting, includes the log of the stochastic error. I allow the errors to be correlated across commuting pairs.

The estimated distance-elasticity of commuting is -0.045 reported in Column (3) of Table 5, which is statistically significant at the 1% significance level.[37] This time, the magnitude of the distance-elasticity is smaller than the estimates available in the literature. Ahlfeldt et al. (2015) estimate 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. Given the large difference in the estimates, 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:

$$\ln \pi_{rm,t} = \tilde{\phi}_{r,t} + \tilde{\phi}_{m,t} - \kappa\epsilon d_{rm} + \varepsilon^{com}_{rm,t}, \tag{18}$$

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 $\kappa\epsilon$,

[37] In Appendix, Table B.5 presents the estimated distance-elasticity of commuting based on Eq. 17 starting without any fixed effects and gradually adding them. In the last column, I show that the distance-elasticity does not vary across the sample periods.

the error term $\varepsilon^{com}_{rm,t}$ 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 Eq. 12 with time subscripts whenever applicable has a direct correspondence with Equation (18). The residual term $\varepsilon^{com}_{rm,t}$ is equal to $\ln AMMA_{rm,t}\varepsilon_{rm,t} = \ln \sum_{o=1}^{J} \frac{\phi_{o,t}}{(\varepsilon_{orm,t} D_{or} D_{om})^{\epsilon}}$. 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 (direct) commuting cost is associated with higher indirect costs of commuting due to migration and job finding, which implies a low value of $AMMA_{rm,t}$. Column (4) reports the estimated semi-elasticity based on Eq. 18.[38] The estimate is -0.074, which is more negative compared to the estimate in Column (3) in line with the intuition and close to the estimate in the literature.

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 (14):

$$\ln \pi_{orm,t} = \phi_{rm,t} + \phi_{or,t} - \delta\epsilon d_{om} + \varepsilon^{jf}_{orm,t}, \tag{19}$$

where the commute-pair fixed effects $\phi_{rm,t}$ capture net benefits of living in r and working in m, $\ln \frac{\phi_{r,t}\phi_{m,t}}{D^{\epsilon}_{rm}}$ (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 $\varepsilon^{jf}_{orm,t}$ captures the random noise. I allow the errors to be correlated across job finding pairs.[39] Column (5) of Table 5 reports the estimated distance-elasticity of job finding. One kilometer increase between origin and workplace location decreases the probability of job finding by 1.57 percentage points. This estimate is both statistically significantly different from zero at 1% significance level and economically important.

5.3 Implications

Taking the gravity framework to the spatial distribution of workers in South Korea, I find that three distinct dimensions of spatial linkages between localities (costs of migration, commuting, and job finding) are important determinants of the spatial distribution of workers. In particular, they

[38] In Appendix, Table B.6 presents the estimated distance-elasticity of job finding based on Eq. 19 starting without any fixed effects and gradually adding them. In the last column, I show that the distance-elasticity does not vary across the sample periods.

[39] 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.

are systematically explained by distance. The estimated reduced-form elasticities are negative and stable over time.[40]

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.[41] I find substantial biases with the estimates of the distance elasticities of migration and commuting reported in the previous literature because of the lack of data on the joint distribution of migration and commuting. The estimated elasticities of migration available in the literature are likely biased toward zero because the cost of migration is positively correlated with the benefits from changing residences ($ALMA$), which are heterogeneous depending on where workers migrate from, due to both commuting and job finding frictions . The available estimates of 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.[42]

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 fiscal policies on taxation, revenue sharing, and the rules of redistribution. In equilibrium, wages, floor space prices, and local government spending are endogenously determined along with the spatial distribution of workers. Lastly, I define the spatial general equilibrium of the economy.

[40]This finding, in particular related to migration, is consistent with the assumption on migration friction in Caliendo et al. (2019), who extend 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.

[41]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.

[42]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 observed spatial distribution of workers. Consistent with these findings, the model presented in this paper allows workers to make migration and commuting decisions jointly.

6.1 More on Worker's Location Decisions

I characterize the market clearing conditions for migration and commuting based on the gravity equation (Eq. 6) 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:

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

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:

$$R_r = \sum_{o=1}^{J} \pi_{r|o} \pi_o R = \sum_{o=1}^{J} \frac{\left(\frac{\tilde{B}_r}{D_{or} Q_r^{1-\beta}} \left(\frac{G_r}{R_r^{\theta}} \right)^{\lambda} \right)^{\epsilon} ALMA_{or}}{\sum_{r'=1}^{J} \left(\frac{\tilde{B}_{r'}}{D_{or'} Q_{r'}^{1-\beta}} \left(\frac{G_{r'}}{R_{r'}^{\theta}} \right)^{\lambda} \right)^{\epsilon} ALMA_{or'}} \pi_o R. \tag{20}$$

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:

$$\pi_{m|r} = \frac{\sum_{o=1}^{J} \pi_{orm}}{\sum_{m'=1}^{J} \sum_{o'=1}^{J} \pi_{o'r'm'}} = \frac{\sum_{r0=1}^{J} \Phi_{orm} \pi_o / \Phi_o}{\sum_{m'=1}^{J} \sum_{o'=1}^{J} \Phi_{o'rm'} \pi_{o'} / \Phi_{o'}}$$
$$= \frac{\left(\frac{(1-\tau_m) \tilde{w}_m}{D_{rm}} \right)^{\epsilon} \sum_{o=1}^{J} \frac{\pi_o / \Phi_o}{(D_{or} D_{om})^{\epsilon}}}{\sum_{m'=1}^{J} \left(\frac{(1-\tau_{m'}) \tilde{w}_{m'}}{D_{rm'}} \right)^{\epsilon} \underbrace{\sum_{o'=1}^{J} \frac{\pi_{o'} / \Phi_{o'}}{(D_{o'r} D_{o'm})^{\epsilon}}}_{AMMA_{rm'}}},$$

where the terms specific to current residence such as amenities, housing prices, and government goods are canceled out from the numerator and denominator. Workers are more likely to commute to places with higher returns $((1-\tau_m)\tilde{w}_m)^\epsilon$ net of commuting costs 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 to enable the commute, $\sum_{o=1}^{J} \frac{\pi_o/\Phi_o}{(D_{or}D_{om})^\epsilon}$, which corresponds to augmented migrant market access $AMMA$. Using these probabilities, I obtain the following expression:

$$L_m = \sum_{r=1}^{J} \pi_{m|r} R_r = \sum_{r=1}^{J} \frac{\left(\frac{(1-\tau_m)\tilde{w}_m}{D_{rm}}\right)^\epsilon AMMA_{rm}}{\sum_{m'=1}^{J} \left(\frac{(1-\tau_{m'})\tilde{w}_{m'}}{D_{rm'}}\right)^\epsilon AMMA_{rm'}} R_r, \tag{21}$$

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:

$$\mathbb{E}[(1-\tau_m)w_m|r] = \sum_{m=1}^{J} \frac{\left(\frac{(1-\tau_m)\tilde{w}_m}{D_{rm}}\right)^\epsilon AMMA_{rm}}{\sum_{m'=1}^{J} \left(\frac{(1-\tau_{m'})\tilde{w}_{m'}}{D_{rm'}}\right)^\epsilon AMMA_{rm'}} (1-\tau_m)w_m. \tag{22}$$

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

$$H_r^R = (1-\beta)\frac{\mathbb{E}[(1-\tau_m)w_m|r]R_r}{Q_r}. \tag{23}$$

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,

$$\mathbb{E}[u_o] = \Gamma(\frac{\epsilon-1}{\epsilon})\Phi_o^{1/\epsilon} = \Gamma(\frac{\epsilon-1}{\epsilon}) \left[\sum_{r'=1}^{J}\sum_{m'=1}^{J} \left(\frac{\tilde{B}_{r'}(1-\tau_{m'})\tilde{w}_{m'}}{D_{or'm'}Q_{r'}^{1-\beta}} \left(\frac{G_{r'}}{R_{r'}^{\theta}}\right)^\lambda\right)^\epsilon\right]^{1/\epsilon} \equiv \bar{u}_o, \tag{24}$$

where the expectation is taken over the distribution of the idiosyncratic component of utility.[43] Based on the expression for the expected utility, I construct a measure of economy-wide welfare while taking into account that the costs of migration and job finding are sunk and workers experience disutility from commuting. This measure corresponds to consumption equivalent worker welfare.

[43]See Appendix D.1 for the derivation of Equation (24).

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:

$$y_m = A_m L_m^{\alpha} \left\{ H_m^F \right\}^{1-\alpha} \tag{25}$$

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:

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

The equilibrium wage equates the labor demand (26) to the labor supply (21) in each location. Similarly, the demand for floor space is given by

$$H_m^F = (\frac{(1-\alpha)A_m}{Q_m})^{\frac{1}{\alpha}} L_m. \tag{27}$$

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., $H_j^F = (1-\vartheta_j)H_j$):

$$(1-\beta)\frac{\mathbb{E}[(1-\tau_m)w_m|r]R_j}{Q_j} = (1-\vartheta_j)H_j. \tag{28}$$

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., $H_j^R = \vartheta_j H_j$):

$$(\frac{(1-\alpha)A_j}{Q_j})^{\frac{1}{1-\alpha}} L_j = \vartheta_j H_j. \tag{29}$$

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 allowing residential and commercial uses (Ahlfeldt et al., 2015; Monte et al., 2018; Tsivanidis, 2019).[44]

[44]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

6.4 National and Local Governments

Consistent with the national fiscal policies discussed in 2.3, I model how local government spending is determined. First, the national government determines a progressive income tax schedule $\tau(w)$, which is increasing in w and uniformly applied to all districts. Without loss of generality, I express $\tau(w_m) = \tau_m$.Local-national revenue sharing implies 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.

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 a share ς_j of the budget allotted for intergovernmental transfers to be delivered to each local government 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 $\{\{\tau_m\}_{m=1}^{J}, \varsigma, \chi, \{\varsigma_j\}_{j=1}^{J}\}$, residential density (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. The budget balancing equation of local government in district j is expressed as follows

$$G_j = \varsigma \underbrace{\sum_{m=1}^{J} \tau_m w_m \pi_{m|j} R_j}_{TR_j} + \varsigma_j (1-\varsigma)\chi \sum_{j'=1}^{J} TR_{j'}, \tag{30}$$

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 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. This redistribution mechanism has features that are structurally similar

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 made to incorporate land use regulations that limit the return to floor space allotted to commercial use as in Ahlfeldt et al. (2015) and Tsivanidis (2019).

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 (20), commuting market clearing (21), labor market clearing (26), floor space market clearing for residential and commercial uses (28 and 29), local government budget balancing equation (30), and population mobility (24).

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 ($\alpha =$.868) 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 Valentinyi and Herrendorf (2008). Second, I set housing expenditure $1-\beta$ equal to .125 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, $\varsigma =$.091 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:

$$\varsigma_j = \frac{IT_j}{\sum_{j'=1}^{J} IT_{j'}}. \tag{31}$$

Table 6: Summary of Model Parameter Estimates

Parameter	Description	Value	Method	Source
α	Labor share	0.868	Calibrated	Economic Census (2015)
$1-\beta$	Housing expenditure share	0.125	Calibrated	Bank of Korea (Q1-Q4 2015)
ϵ	Fréchet shape parameter	3.429	Estimated	Fréchet Property
λ	Preference for local gov't spending	0.301	Estimated	Gravity Equation
θ	Net congestion	0.780	Estimated	Gravity Equation
ρ	Spatial decay of migration	0.010	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	0.261	Estimated	GE Condition
ς	Local-national revenue Sharing	0.091	Observed	Local Subsidy Act
χ	Extent of Redistribution	0.350	Observed	Local Subsidy Act
$\varsigma_j \forall j$	Redistribution		Recovered	Local Gov't Budget Balance
$A_j \forall j$	Productivity		Recovered	Profit Maximization + Zero Profit
$\tilde{B}_j \forall j$	Adjusted amenities		Recovered	Spatial Mobility
$H_j \forall j$	Stock of 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 from Bank of Korea. 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.

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. Instead, I calibrate a single tax rate τ by estimating the following equation:

$$\ln TR_j = \ln \tau + \ln w_j^R + \xi_j, \tag{32}$$

where TR_j is the total tax revenue paid by workers living in j and w_j^R denotes the total wage earned by the workers (i.e., $w_j^R = w_m \pi_{m|j} R_j$). The estimated value of τ is .261.

7.3 Recovery of Unobserved Local Characteristics

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:

$$Q_j = (1-\alpha)\left(\frac{\alpha}{w_j}\right)^{\frac{\alpha}{1-\alpha}} A_j^{\frac{1}{1-\alpha}}. \tag{33}$$

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 (33). Figure C.1 in Appendix plots the spatial distribution of the recovered values of local productivity in Panel (a). 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).[45]

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:

$$\pi_{m|ro} = \frac{\pi_{orm}}{\sum_{m'=1}^{J} \pi_{orm'}} = \frac{\Phi_{orm}}{\sum_{m'=1}^{J} \Phi_{orm'}} = \frac{\frac{\tilde{w}_m^{\epsilon}}{\exp(\kappa\epsilon d_{rm}+\delta\epsilon d_{om})}}{\sum_{m'=1}^{J} \frac{\tilde{w}_{m'}^{\epsilon}}{\exp(\kappa\epsilon d_{rm'}+\delta\epsilon d_{om'})}}. \tag{34}$$

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 $\kappa\epsilon$ and $\delta\epsilon$ and rearranging such that left hand side consists of only observables, I obtain the following expression:

$$\ln \pi_{m|ro} + \kappa\epsilon d_{rm} + \delta\epsilon d_{om} = -\ln \sum_{m'=1}^{J} \frac{\omega_{m'}}{\exp(\kappa\epsilon d_{rm'} + \delta\epsilon d_{om'})} + \ln \omega_m, \tag{35}$$

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 (35), 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.

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 3.43; this means the worker mobility increases by 3.43 percent for a 1 percent increase in wages.

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

[45] The binned scatter plot in Panel (b) of Figure C.1 shows that the districts with higher levels of productivity have a greater number of firms.

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 ($\lambda = \tilde{\lambda}/\epsilon$ =0.301, $\rho = \tilde{\rho}/\epsilon$ =0.010, $\kappa = \tilde{\kappa}/\epsilon$ =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 ϵ =3.43, the implied value of $1-\beta$ is equal to 0.143, close to the expenditure share estimated using the Household Expenditure Survey in Section 7.1 ($1-\beta$ =.125.

Based on the structural value of how much people of local government spending (λ =0.301), 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.

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 (6):

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

I take the log transformation of both sides of Equation (36) and rearrange such that left hand side only consists of observables:

$$\ln \pi_{r|mo} - \ln \frac{G_r^{\lambda\epsilon}}{\exp(\kappa\epsilon d_{rm} + \rho\epsilon d_{or}) Q_r^{(1-\beta)\epsilon} R_r^{\theta\lambda\epsilon}} = -\ln \sum_{r'=1}^{J} \frac{\tilde{B}_{r'}^{\epsilon} G_{r'}^{\lambda\epsilon}}{\exp(\kappa\epsilon d_{r'm} + \rho\epsilon d_{or'}) Q_{r'}^{(1-\beta)\epsilon} R_{r'}^{\theta\lambda\epsilon}} + \ln \tilde{B}_r^{\epsilon}, \tag{37}$$

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.2 in Appendix plots the spatial distribution of adjusted amenities in Panel (a). 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.[46]

[46] I assess the correlation between the recovered amenities and a local outcome which ma.In Panel (b) of Figure C.2, residences with lower suicide rates tend to have higher amenities. 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

7.4 Non-targeted Moments

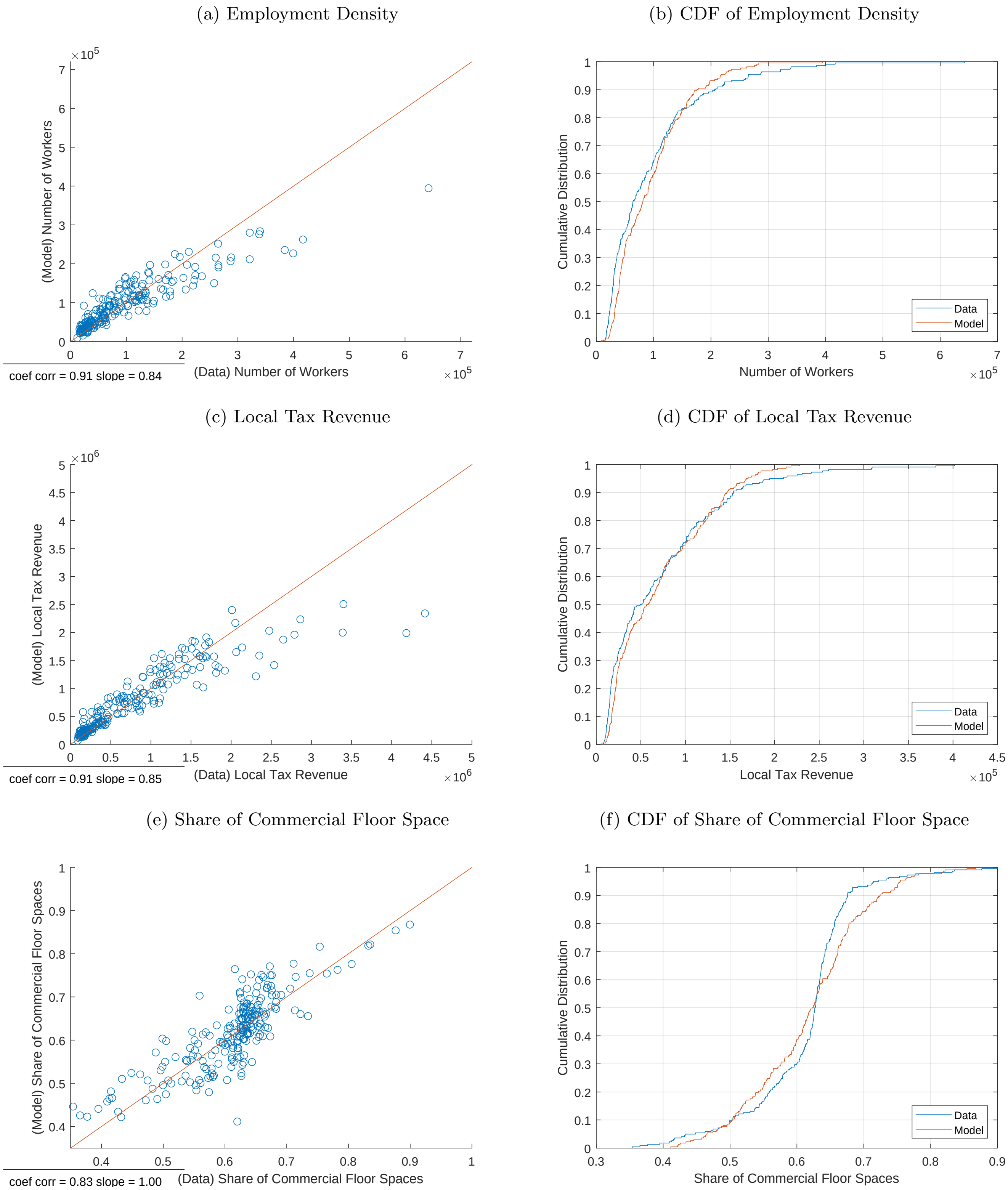
I evaluate how well the model predicts the non-targeted moments. First, I compare the observed data on number of workers by employment location to the model prediction in Panel (a) and (b) of Figure 4. The two variables have a coefficient correlation of 0.94 with a slope equal to 0.91 in Panel (a). The 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 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 for residential purposes, which does not take into account the ratio of floor space to land area. Despite the sources of measurement error, the model performs well in capturing residential floor spaces.

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.

Figure 4: Over-identifying Moments: Model vs. Data



Note: This figure compares 2015 data with the model predictions of non-targeted moments. Panel (a) and (b) plot the 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.

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 B.1 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.

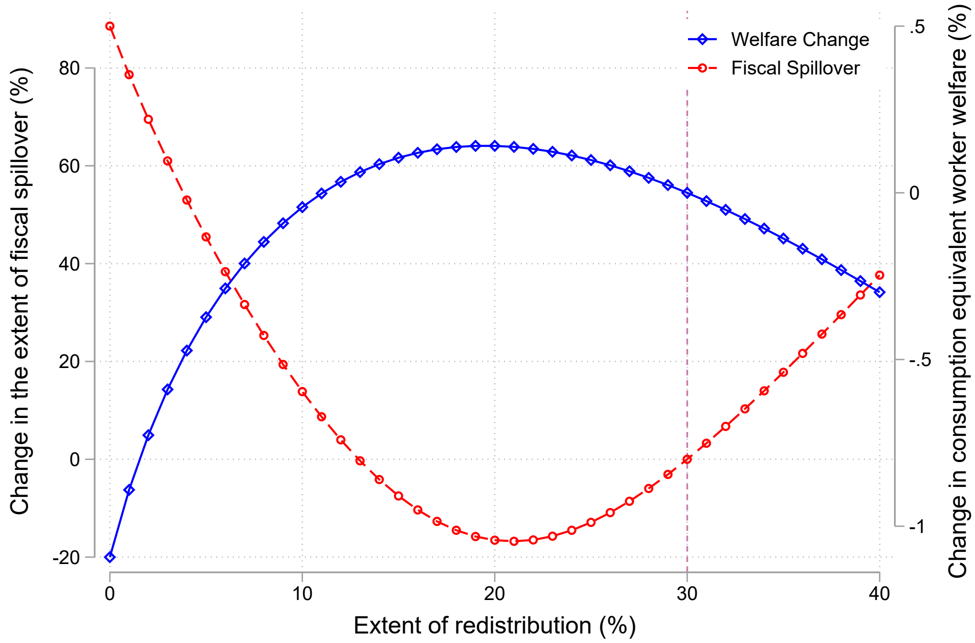
8.2 Welfare Consequences of Redistribution

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:

$$G_j = (\tilde{\chi} - \tilde{\varsigma})TR_j + \varsigma_j \tilde{\varsigma} \sum_{j'=1}^{S} TR_{j'} \tag{38}$$

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 5: Spatial Fiscal Redistribution and Aggregate Welfare Changes



Note: 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. 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 5 plots the changes in the aggregate welfare of workers 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 1 percentage point), 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.[47] Therefore, in the presence of redistributive intergovernmental transfers,

[47] 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

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

Therefore, the welfare changes summarized in Figure 5 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.

In addition, Figure 5 plots 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 zero 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 varying extents of redistribution.

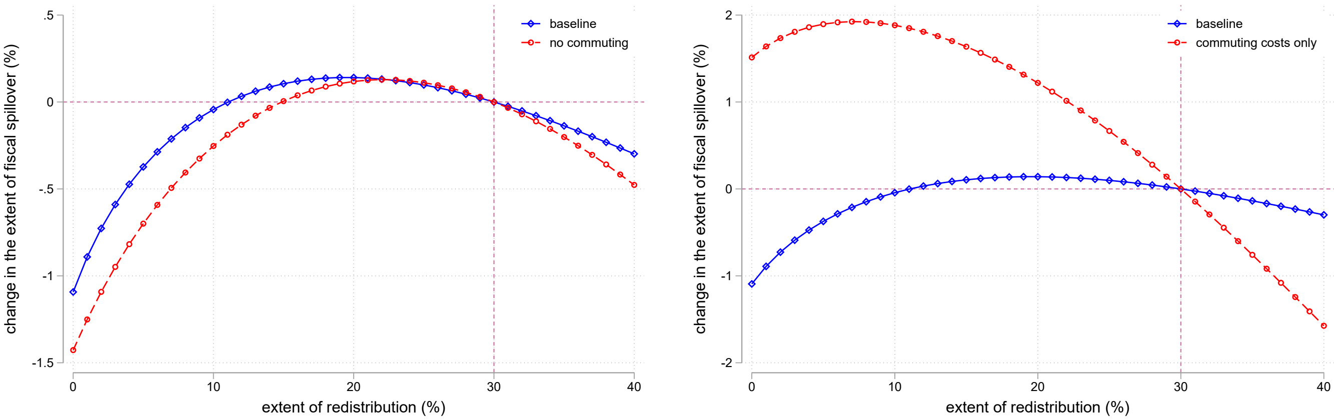
In Panel (a) of Figure 6, I plot the welfare changes relative to the baseline in 2015 assuming no inter-district commuting. If workers are not allowed to commute outside of districts, eliminating redistribution altogether leads in a higher welfare loss of about 2 percent. Furthermore, the optimal extent 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 6 plots the changes in the worker welfare under the assumption of no

diverted for redistribution.

Figure 6: Aggregate Welfare Changes under Alternative Assumptions in the Literature

(a) Prohibitively Costly Commuting (b) Zero Cost of Migration and Job Finding



Note: 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 migration literature and assume workers cannot work outside of their district of residence. I solve for a new equilibrium for 2015 assuming the distance elasticities of migration equal to 0.007 as in Column (2) of Table 5, 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 percent) in Panel (a). Second, I follow the assumption commonly imposed in the commuting literature: i.e., there is no costs of migration 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 (4) 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).

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

The observed uneven distribution of local public spending mirrors the uneven spatial distribution of economic activity and may become more or less unequal depending on the extent of spatial redistribution of tax revenue. In this paper, I examined the aggregate welfare consequences of spatial fiscal redistribution through the lens of a quantitative spatial general equilibrium model in the case of South Korea. The model simultaneously features two margins of mobility—migration and commuting—that have been previously studied separately. I combine the framework with the quasi-natural experiment as a source of exogenous variation and the unique data capturing the joint distribution of migration and commuting to estimate the key reduced-form elasticities of worker mobility with respect to local government expenditure, residential density, and home prices. I conducted a conuterfactual policy experiment in which I vary how much redistributive transfers contribute to local public spending to shed light on the optimal extent of redistribution. At the optimal mix of local taxation and redistributive intergovernmental transfers, the spatial variation in real local public spending is minimized as the inter-regional fiscal spillovers generated by redistribution counteract the intra-regional spillovers. Throughout the paper, I show that it is crucial to account for both margins of mobility—migration and commuting. Studying one margin alone not only biases the estimates of key structural parameters we care about in the urban, public, and spatial literature, but also may produce erroneous welfare implications of fiscal spatial redistribution.

References

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015). The economics of density: Evidence from the berlin wall. *Econometrica 83*(6), 2127–2189.

Akcigit, U., S. Baslandze, and S. Stantcheva (2016). Taxation and the international mobility of inventors. *American Economic Review 106*(10), 2930–2981.

Albert, C. and J. Monras (2019). Immigration and spatial equilibrium: the role of expenditures in the country of origin. *Working Paper*, 1–86.

Albouy, D. (2012). Evaluating the efficiency and equity of federal fiscal equalization. *Journal of Public Economics 96*, 824–839.

Almagro, M. and T. Domínguez-Iino. Location sorting and endogenous amenities: Evidence from Amsterdam.

Balboni, C. (2019). In harm's way? infrastructure investments and the persistence of coastal cities. *Working Paper*, 1–86.

Bertrand, M., E. Duflo, and S. Mullainathan (2004, February). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics 119*(1), 249–275.

Bianchi, N., M. Giorcelli, and E. M. Martino (2019). The effects of fiscal decentralization on publicly provided services and labor markets. *Working Paper*, 1–40.

Black, S. E. (1999, May). Do better schools matter? parental valuation of elementary education. *The Quarterly Journal of Economics 114*(2), 577–599.

Bryan, G. and M. Morten (2019). The aggregate productivity effects of internal migration: Evidence from indonesia. *Journal of Political Economy 127*(5), 2229–2268.

Cadena, B. and B. Kovak (2016). Immigrants equilibrate local labor markets: Evidence from the great recession. *American Economic Journal: Applied Economics 8*(1), 257–290.

Caliendo, L., M. Dvorkin, and F. Parro (2019, May). Trade and labor market dynamics: General equilibrium analysis of the china trade shock. *Econometrica 87*(3), 741–835.

Card, D. (2001). Immigration inflows, native outflows, and the local market impacts of higher immigration. *Journal of Labor Economics 19*(1), 22–64.

Cellini, S. R., F. Ferreira, and J. Rothstein (2010). The value of school facility investments: Evidence from a dynamic regression discontinuity design. *The Quarterly Journal of Economics 125*(1), 215–261.

Chay, K. Y. and M. Greenstone (2005, April). Does air quality matter? evidence from the housing market. *Journal of Political Economy 113*(2), 376–424.

Ciccone, A. and R. Hall (1996). Productivity and the density of economic activity. *American Economic Review 86*(1), 54–70.

Combes, P. and L. Gobillon (2015). The empirics of agglomeraion economics. In G. Duranton, V. Henderson, and W. Strange (Eds.), *Handbook of Regional and Urban Economics*, Chapter 5, pp. 247–348. Amsterdam: Elsevier.

Davis, M. A. and F. Ortalo-Magné (2011). Household expenditures, wages, rents. *Review of Economic Dynamics 14*, 248–261.

de la Roca, J. and D. Puga (2017). Learning by working in big cities. *Review of Economic Studies 84*, 106–142.

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

Dix-Carneiro, R. (2014, May). Trade liberalization and labor market dynamics. *Econometrica 82*(3), 825–885.

Donaldson, D. and R. Hornbeck (2016). Railroads and american economic growth: A market access approach. *Quarterly Journal of Economics 131*(2), 799–858.

Ducruet, C., R. Juhász, D. K. Nagy, and C. Steinwender (2019). All aboard: The aggregate effects of port development. *Working Paper*, 1–70.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70*(5), 1741–1779.

Eeckhout, J., R. Pinheiro, and K. Schmidheiny (2014). Spatial sorting. *Journal of Political Economy 122*(3), 554–620.

Fajgelbaum, P. D. and C. Gaubert (2018). Optimal spatial policies, geography and sorting. *NBER Working Paper 24632*, 1–64.

Fajgelbaum, P. D., E. Morales, J. C. Suarez-Serrato, and O. Zidar (2019). State taxes and spatial misallocation. *Review of Economic Studies 86*(1), 333–376.

Fisman, R. and R. Gatti (2002). Decentralization and corruption: Evidence from countries. *Journal of Public Economics 83*, 325–345.

Gelbach, J. B. (2004). Migration, the life cycle, and state benefits: How low is the bottom? *Journal of Political Economy 112*(5), 1091–1130.

Glaeser, E. and J. Gottlieb (2008). The economics of place-making policies. *Brookings Papers on Economic Activity Spring*, 155–239.

Glaeser, E. and D. Maré (2001). Cities and skills. *Journal of Labor Economics 19*(2), 316–342.

Kleven, H. J., C. Landais, E. Saez, and E. Schultz (2014). Migration and wage effects of taxing top earners: Evidence from the foreigners' tax scheme in denmark. *The Quarterly Journal of Economics*, 333–378.

Kline, P. and E. Moretti (2014). People, places, and public policy: Some simple welfare economics of local economic development programs. *The Annual Review of Economics 6*, 629–662.

Lee, J. (2017). The labor market in south korea, 2000-2016. Technical report, IZA.

Manning, A. and B. Petrongolo (2017). How local are labor markets? evidence from a spatial job search model. *American Economic Review 107*(10), 2877–2907.

Marinescu, I. and R. Rathelot (2018, July). Mismatch unemployment and the geography of job search. *American Economic Journal: Macroeconomics 10*(3), 42–70.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics 3*(4), 303–328.

Molloy, R., C. L. Smith, and A. Wozniak (2011). Internal migration in the united states. *Journal of Economic Perspective 25*(3), 173–196.

Monte, F., S. Redding, and E. Rossi-Hansberg (2018). Commuting, migration, and local employment elasticities. *American Economic Review 108*(12), 2855–3890.

Moretti, E. and D. J. Wilson (2017). The effect of state taxes on the geographical location of top earners: Evidence from star scientists. *American Economic Review 107*(7), 1858–1903.

Morten, M. and J. Oliveira (2018). The effects of roads on trade and migration: Evidence from a planned capital city. *NBER Working Paper 22158*, 1–64.

Moulton, B. R. (1990, May). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The Review of Economics and Statistics 72*(2), 334–338.

Oates, W. E. (1999). An essay on fiscal federalism. *Journal of Economic Literature 37*(3), 1120–1149.

OECD (2016). Better life index: Korea. Technical report.

Pellegrina, H. S. and S. Sotelo (2019, June). Migration, specialization, and trade: Evidence from the brazilian march to the west. *Working Paper*, 1–43.

Piyapromdee, S. (2017). The impact of immigration on wages, internal migration, and welfare. *Working Paper*, 1–83.

Preston, S., P. Heuveline, and M. Guillot (2000, October). *Demography: Measuring and Modeling Population Processes*. Wiley-Blackwell.

Rappaport, J. (2007). Moving to nice weather. *Regional Science and Urban Economics 37*, 375–398.

Redding, S. and E. Rossi-Hansberg (2017). Quantitative spatial economics. *Annual Review of Economics 9*, 21–58.

Redding, S. and M. Turner (2015). Transportation costs and the spatial orgnization of economic activity. In G. Duranton, V. Henderson, and W. Strange (Eds.), *Handbook of Regional and Urban Economics*, Chapter 20, pp. 1339–1398. Amsterdam: Elsevier.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of Political Economy 90*(6), 1257–1278.

Rosen, S. (1979). Wages-based indexes of urban quality of life. *Current Issues in Urban Economics*, 34–55.

Sanderson, E. and F. Windmeijer (2016). A weak instrument f-test in linear iv models with multiple endogenous variables. *Journal of Econometrics 190*, 212–221.

Schafer, A. (2000). Regularities in travel demand: an international perspective. *Journal of Transportation Statistics 3*, 1–31.

Stock, J., J. Wright, and M. Yogo (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics 20*(4), 518–529.

Suárez-Serrato, J. C. and P. Wingender (2014). Estimating the incidence of government spending. *Working Paper*, 1–69.

Suárez-Serrato, J. C. and P. Wingender (2016). Estimating local fiscal multipliers. *Working Paper*, 1–99.

Tiebout, C. (1956). A pure theory of local expenditures. *Journal of Political Economy 64*(5), 416–424.

Tsivanidis, N. (2019). Evaluating the impact of urban transit infrastructure: Evidence from Bogotá's Trans-Milenio. *Working Paper*, 1–56.

Valentinyi, A. and B. Herrendorf (2008). Measuring factor income shares at the sectoral level. *Review of Economic Dynamics 11*, 820–835.

Wildasin, D. (1980, November). Locational efficiency in a federal system. *Regional Science and Urban Economics 10*, 453–471.

The Spatial Distribution of Public Spending, Commuting, and Migration

Online Appendix

Wookun Kim

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.

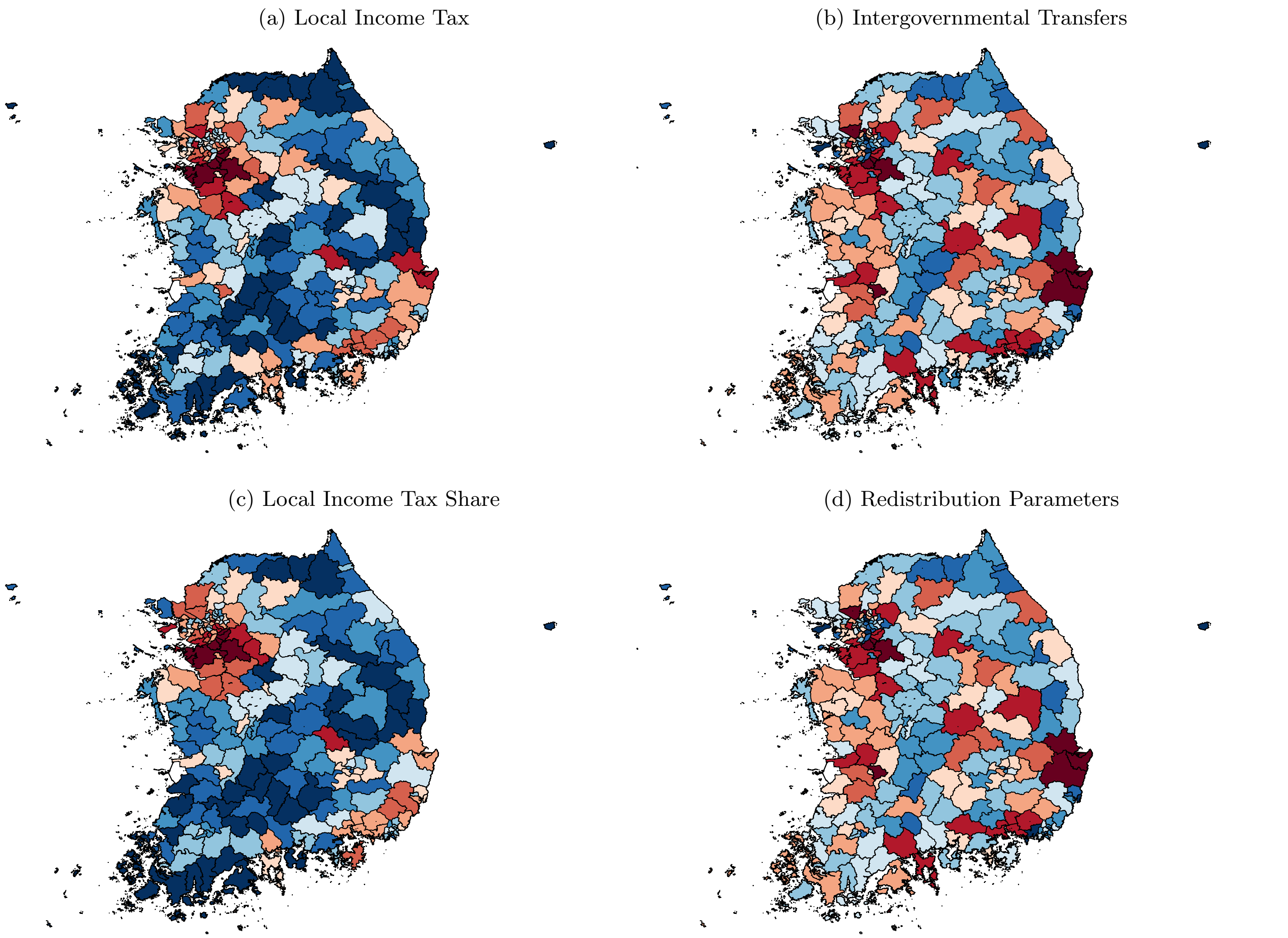
Annual Migration Rates

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. Notwithstanding 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

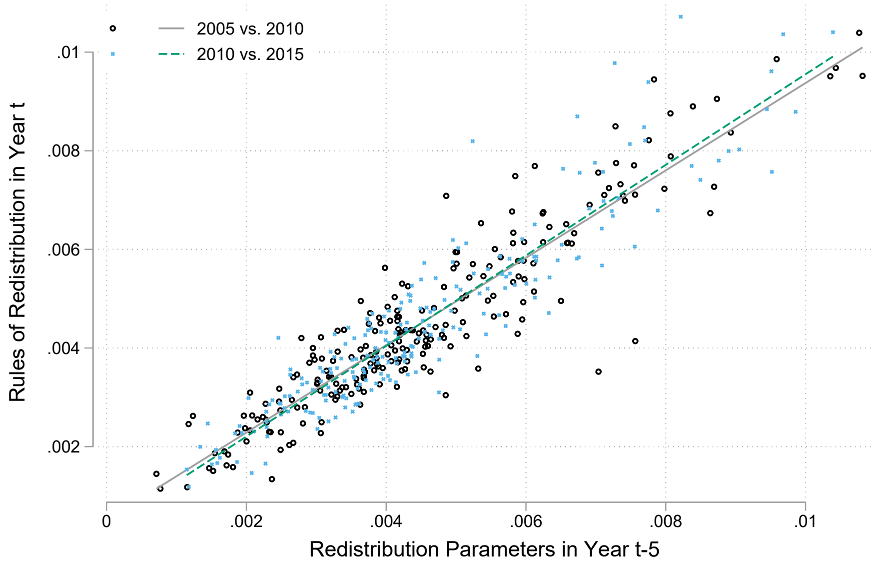
B.1 Local Income Taxes and Intergovernmental Transfers

Figure B.1: Sources of Local Government Spending



Note: This figure describes the spatial distribution of local government spending by its revenue sources observed in 2015. Panel (a) and (b) plot the spatial distribution of local government revenue by its sources: local income taxes and intergovernmental transfers. Panel (c) shows the contribution of local income tax to local government spending by districts. Panel (d) plots the rules of redistribution equal to the amount of intergovernment transfers redistributed to each district divided by the total amount of intergovernment transfers across all districts.

Figure B.2: Redistribution Parameters over Time



Note: this figure compares the rules of redistribution across year 2005, 2010, and 2015. For each year, the rules of redistribution is a set of district-specific policy parameter values equal to the amount of intergovernmental transfers each district receives divided by the total amount of intergovernmental transfers rebated back to all districts. The estimated slopes are 0.982 (0.012) for 2005 vs. 2010 and 0.988 (0.012) for 2010 vs. 2015.

B.2 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 (7), 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 (15), at the commuting-pair level when estimating Equation (17), and at the job-finding-pair level when estimating Equation (19).

Table B.1: Determinants of Redistribution Policy

	(1)	(2)	(3)	(4)
Dependent Variable:	Rules of Redistribution ($\ln \varsigma_r$)			
Residential Population ($\ln R_r$)	0.3730***	0.4071***	0.4595***	0.5251***
	(0.0170)	(0.0183)	(0.0183)	(0.0585)
Residential Amenities ($\ln \tilde{B}_r$)		−0.1242***	−0.1321***	−0.1324***
		(0.0264)	(0.0212)	(0.0211)
Workplace Productivity ($\ln A_r$)			−0.6222***	−0.5427***
			(0.0781)	(0.1114)
Employment Population ($\ln L_r$)				−0.0795
				(0.0694)
Area ($\ln Area_r$)	0.2023***	0.1962***	0.1822***	0.1848***
	(0.0109)	(0.0113)	(0.0090)	(0.0085)
Observations	222	222	222	222
Adjusted R^2	0.663	0.698	0.774	0.774

Notes: In this table, I investigate the determinants of the rules of redistribution by projecting them on local charateristics. The dependent variable is the log of the share of intergovernmental transfers received in 2015. I begin with covariates of residential density and area in Column (1) and gradually introduce additional variables across columns. Each observation corresponds to a district in 2015.

B.3 Travel Time vs. Distance of Commuting

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

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

The results are presented in Table 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)

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

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Commuting Time ($\tau^{(time)}_{rm,t}$)				
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: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

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.4 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 (7) would include the distribution of workers by education $\pi_{edu|r,t}$.

The exclusion restriction (9) 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:

$$\pi_{b|r,t} = \phi_r + \eta_{b',b}\tau_{b',t} + \zeta_{b,r,t}, \tag{A.1}$$

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.3: Tax reforms did not affect education distribution

	(1)	(2)	(3)
A. Educational Attainment: Low ($\pi_{low\|r,t}$)			
Tax Rate (Low) $\tau_{low,t}$	1.37e-10		-1.64e-9
	(0.00115)		(0.00169)
Tax Rate (High) $\tau_{high,t}$		1.62e-10	1.19e-9
		(0.00074)	(0.00104)
B. Educational Attainment: High ($\pi_{high\|r,t}$)			
Tax Rate (Low) $\tau_{low,t}$	-1.55e-09		-1.50e-08
	(0.00058)		(0.00134)
Tax Rate (High) $\tau_{high,t}$		-3.88e-10	8.96e-09
		(0.00038)	(.000904)
Observations	666	666	666

Notes: This table reports the estimation results based on Equation (A.1). 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: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table B.3 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.5 Distance Elasticities of Migration, Commuting, and Job Finding

Table B.4: Distance Elasticity of Migration

	(1)	(2)	(3)	(4)
Dependent Variable:	$\ln \pi_{orm,t}$	$\ln \pi_{orm,t}$	$\ln \pi_{orm,t}$	$\ln \pi_{orm,t}$
Distance, d_{or} $(-\rho\epsilon)$	-0.0026^{***}	-0.0044^{***}	-0.0328^{***}	-0.0328^{***}
	(0.0001)	(0.0001)	(0.0009)	(0.0013)
Distance $\times$ 2005				0.0003
				(0.0019)
Distance $\times$ 2010				-0.0004
				(0.0019)
Observations	257,583	257,583	257,583	257,583
Fixed Effects:				
Commuting Pair $\times$ Year $(\phi_{rm,t})$	N	Y	Y	Y
Job Finding Pair $\times$ Year $(\phi_{om,t})$	N	N	Y	Y

Notes: In this table, I estimate the semi-elasticity of migration with respect to distance based on Eq. 15, starting with a simple estimate without any fixed effects in Column (1) and gradually adding the fixed effects. Column (3) corresponds to based on Eq. 15. 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: $***p < 0.01$, $**p < 0.5$, $*p < 0.1$.

In Table B.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 -0.0026, 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. 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 -0.0044, 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 (15). The estimated semi-elasticity is -0.0328 and means that the probability of migration decreases by 3.28 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 toward zero rising from failing to account for the difficulty in finding jobs for workers who migrate 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.

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 Eq. 15: 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.[48]

Table B.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 -0.0083, 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.[49] 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.

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

[49] Travel time is also widely used to define a cost of geographical mobility (Ahlfeldt et al., 2015; Morten and Oliveira, 2018). In Appendix , I show that travel time associated with commuting has a one-to-one relationship with distance of commuting.

Table B.5: Distance Elasticity of Commuting

	(1)	(2)	(3)	(4)
Dependent Variable:	$\ln \pi_{orm,t}$	$\ln \pi_{orm,t}$	$\ln \pi_{orm,t}$	$\ln \pi_{orm,t}$
Distance, d_{rm} $(-\kappa\epsilon)$	-0.0083^{***}	-0.0350^{***}	-0.0450^{***}	-0.0460^{***}
	(0.0004)	(0.0005)	(0.0006)	(0.0009)
Distance $\times$ 2005				0.0008
				(0.0013)
Distance $\times$ 2010				0.0028^{**}
				(0.0013)
Observations	257,583	257,583	257,583	257,583
Fixed Effects:				
Migration Pair $\times$ Year ($\phi_{or,t}$)	N	Y	Y	Y
Job Finding Pair $\times$ Year ($\phi_{om,t}$)	N	N	Y	Y

Notes: In this table, I estimate the semi-elasticity of commuting with respect to distance based on Eq. 17, starting with a simple estimate without any fixed effects in Column (1) and gradually adding the fixed effects. Column (3) corresponds to based on Eq. 17. 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$\times$year, commuting pair$\times$year, job finding pair$\times$year, and a triplet of previous and current residences and workplace: $***p < 0.01$, $**p < 0.5$, $*p < 0.1$.

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 B.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 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 Eq. 19. 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

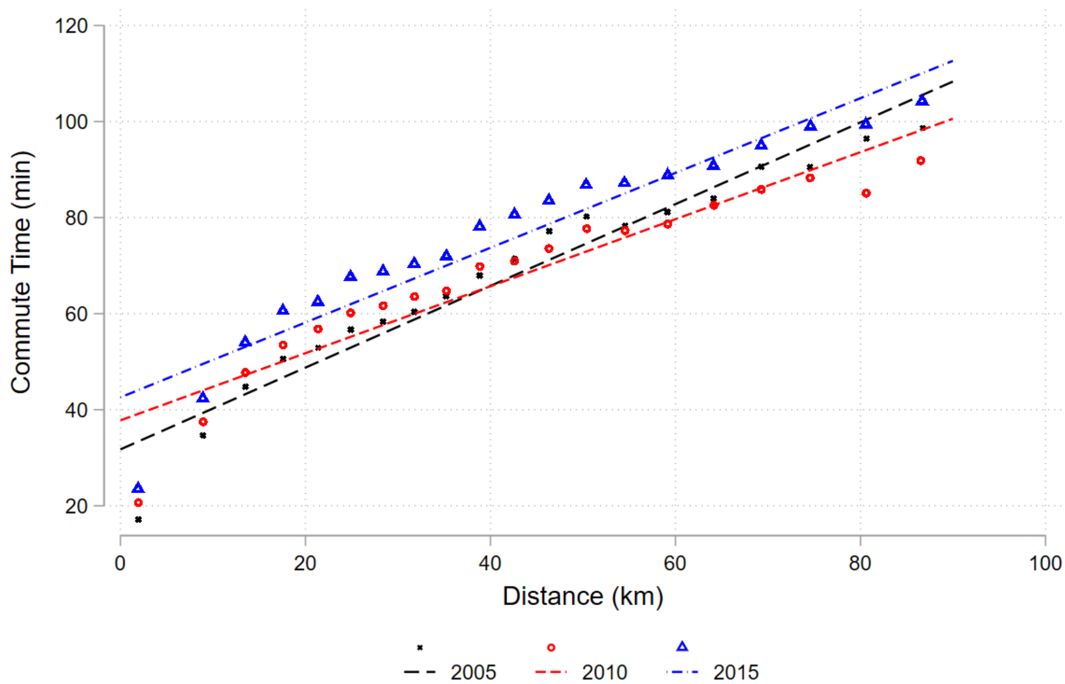
Table B.6: Distance Elasticity of Commuting

	(1)	(2)	(3)	(4)
Dependent Variable:	$\ln \pi_{orm,t}$	$\ln \pi_{orm,t}$	$\ln \pi_{orm,t}$	$\ln \pi_{orm,t}$
Distance, d_{om} $(-\delta\epsilon)$	-0.0024^{***}	-0.0040^{***}	-0.0157^{***}	-0.0150^{***}
	(0.0001)	(0.0001)	(0.0003)	(0.0004)
Distance $\times$ 2005				-0.0015^{**}
				(0.0007)
Distance $\times$ 2010				-0.0009
				(0.0006)
Observations	257,583	257,583	257,583	257,583
Fixed Effects:				
Commuting Pair $\times$ Year ($\phi_{rm,t}$)	N	Y	Y	Y
Migration Pair $\times$ Year ($\phi_{or,t}$)	N	N	Y	Y

Notes: In this table, I estimate the semi-elasticity of job finding with respect to distance based on Eq. 19, starting with a simple estimate without any fixed effects in Column (1) and gradually adding the fixed effects. Column (3) corresponds to based on Eq. 19. 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$\times$year, commuting pair$\times$year, job finding pair$\times$year, and a triplet of previous and current residences and workplace: $***p < 0.01$, $**p < 0.5$, $*p < 0.1$.

between the estimate in Column (4) and the estimate reported in Column (3) is economically small and is not statistically significant.

Figure B.3: Travel Time vs. Distance of Commuting



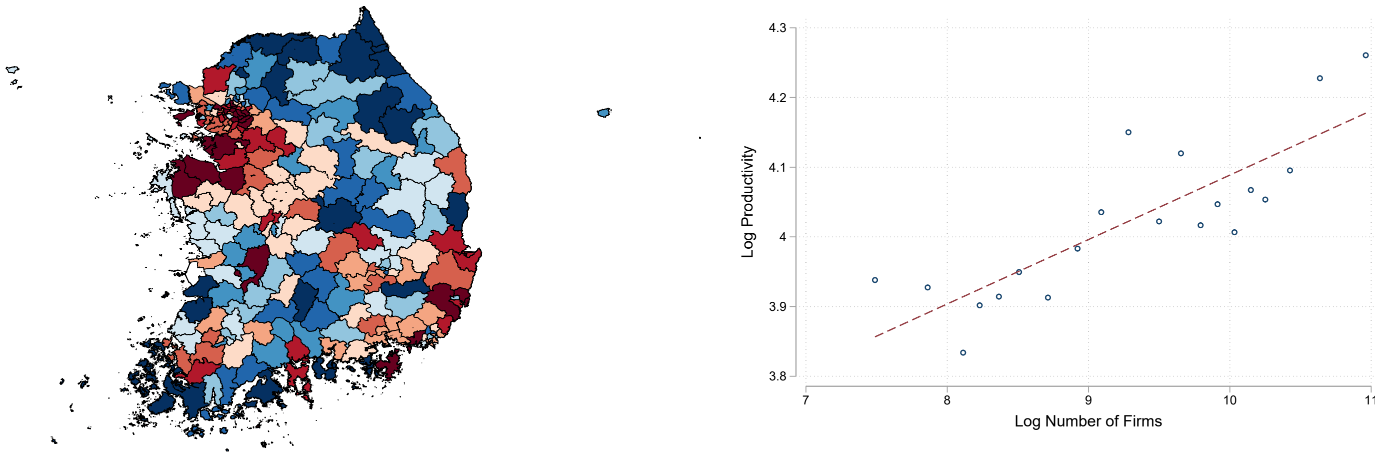
Notes: This figure plots average commuting time in minutes for each of 5 percentiles of commuting distance for each survey year (2005, 2010, and 2015) of the Population Census of South Korea.

C Supplementary Quantitative Results

C.1 Local Productivity

Figure C.1: Recovered Values of Local Productivity Fundamentals

(a) Spatial Distribution (b) Productivity vs. Number of Firms

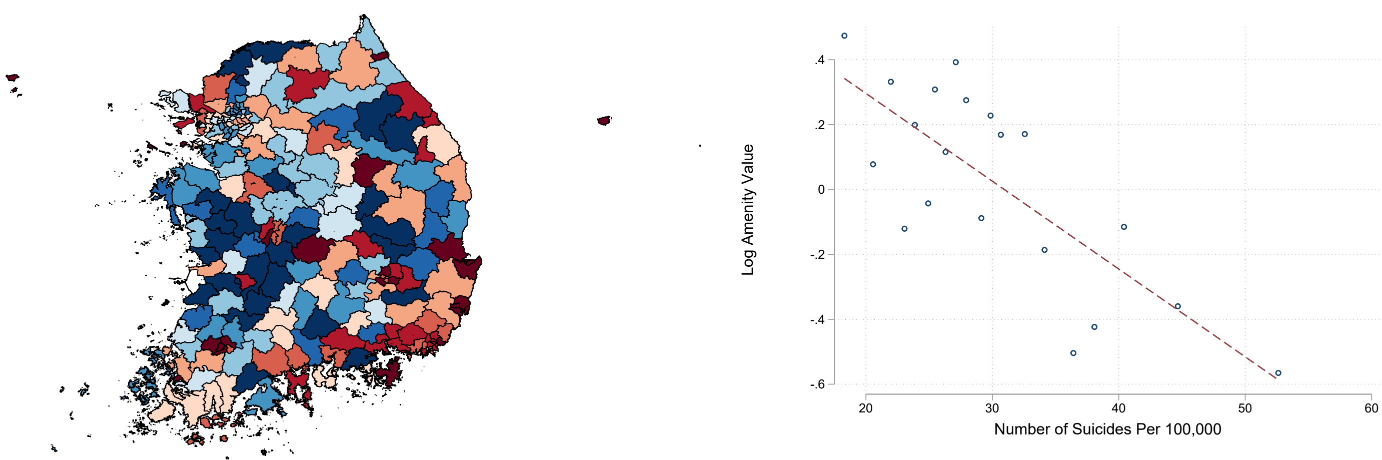


Note: This figure describes the spatial distribution of local productivity fundamentals recovered using the model and the data in 2015 in Panel (a) and its correlation with the number of firms in 2015 in Panel (b).

C.2 Adjusted Amenities

Figure C.2: Recovered Values of Local Amenity Fundamentals

(a) Spatial Distribution (b) Amenity Values vs. Suicide Rates



Note: This figure describes the spatial distribution of local amenity fundamentals recovered using the model and the data in 2015 in Panel (a) and its correlation with suicide rates in 2015 in Panel (b).

C.3 Fiscal Decentralization Policy Parameters

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

- Local government spending:

$$G_r = \varsigma \sum_{m=1}^{S} \tau_m w_m \pi_{m|r} R_r + \varsigma_r (1-\varsigma)\chi \sum_{r'=1}^{S}\sum_{m'=1}^{S} \tau_{m'} w_{m'} \pi_{m'|r'} R_{r'} \tag{A.2}$$

  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 spent 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 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.

C.4 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.4 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 (28). Second, the commercial floor space demanded is determined by local productivity (A_j), employment population (L_j), and floor space prices (Q_j) given labor share in production (α) as in Equation (29). 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.

D entary 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

$$F_{rm}(u) = \Pr[U \leq u] = \Pr(z \leq u \times v_{orm}^{-1}), \tag{A.3}$$

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

$$F_{rm}(u) = \exp(-\frac{T_r M_m B_r (1-\tau_m) w_m}{D_{orm} Q_r^{1-\beta}} \left(\frac{G_r}{R_r^{\theta}}\right)^{\lambda} u^{-\epsilon}) \equiv \exp(-\Phi_{orm} u^{-\epsilon}). \tag{A.4}$$

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:

$$\begin{aligned}
\pi_{rm|o} &= \Pr[u_{rm|o} \geq \max\{u_{jk}\}; \forall j, k] \\
&= \int_0^{\infty} \prod_{k \neq j} F_{rk}(u) \times \left(\prod_{j \neq r} \prod_{k} F_{jk}(u) \right) f_{rm}(u) du \\
&= \int_0^{\infty} \prod_{j} \prod_{k} \epsilon \Phi_{orm} u^{-(\epsilon+1)} \exp(-\Phi_{ojk} u^{-\epsilon}) du \\
&= \int_0^{\infty} \epsilon \Phi_{orm} u^{-(\epsilon+1)} \exp(-\Phi_o u^{-\epsilon}) du,
\end{aligned}$$

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:

$$\pi_{rm|o} = \frac{T_r M_m \left(\frac{B_r (1-\tau_m) w_m}{D_{orm} Q_r^{1-\beta}} \left(\frac{G_r}{R_r^{\theta}} \right)^{\lambda} \right)^{\epsilon}}{\sum_{r'=1}^{J} \sum_{m'=1}^{J} T_{r'} M_{m'} \left(\frac{B_{r'} (1-\tau_{m'}) w_{m'}}{D_{or'm'} Q_{r'}^{1-\beta}} \left(\frac{G_{r'}}{R_{r'}^{\theta}} \right)^{\lambda} \right)^{\epsilon}} \equiv \frac{\Phi_{orm}}{\Phi_o} \tag{A.5}$$

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

$$F_o(u) = \exp(-\Phi_o u^{-\epsilon}), \text{ where } \Phi_o = \sum_{r=1}^{J} \sum_{m=1}^{J} \frac{T_r M_m B_r (1-\tau_m) w_m}{D_{orm} Q_r^{1-\beta}} \left(\frac{G_r}{R_r^{\theta}} \right)^{\lambda}. \tag{A.6}$$

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

$$\mathbb{E}[u|o] = \int_0^\infty \epsilon \Phi_o u^{-\epsilon} e^{-\Phi_o u^{-\epsilon}} du = \Gamma\left(\frac{\epsilon - 1}{\epsilon}\right) \Phi_o^{1/\epsilon} \equiv \bar{u}_o. \tag{A.7}$$

D.2 Isomorphism of the Gravity Equation

I show that the gravity equation (6) is isomorphic to the types of gravity equations derived in the literature on costly movements of people: commuting and migration.

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 **??**. Then, the distribution of workers by current residence and workplace is independent to the distribution of workers by initial residence. Therefore, Equation (6) does not vary by initial residence o and is given by:

$$\pi_{rm} = \frac{T_r M_m \left(\frac{B_r(1-\tau_m)w_m}{D_{rm}Q_r^{1-\beta}}\left(\frac{G_r}{R_r^\theta}\right)^\lambda\right)^\epsilon}{\sum_{r'=1}^S \sum_{m'=1}^S T_{r'} M_{m'} \left(\frac{B_{r'}(1-\tau_{m'})w_{m'}}{D_{r'm'}Q_{r'}^{1-\beta}}\left(\frac{G_{r'}}{R_{r'}^\theta}\right)^\lambda\right)^\epsilon}, \tag{A.8}$$

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., $\lambda = 0$), Eq. A.8 is identical to the gravity equations based on the spatial models of Ahlfeldt et al. (2015) and Monte et al. (2018).

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:

$$\pi_{or} = \frac{T_r M_r \left(\frac{B_r(1-\tau_r)w_r}{D_{or}Q_r^{1-\beta}}\left(\frac{G_r}{R_r^\theta}\right)^\lambda\right)^\epsilon \pi_o}{\sum_{r'=1}^S T_{r'} M_{r'} \left(\frac{B_{r'}(1-\tau_{r'})w_{r'}}{D_{or'}Q_{r'}^{1-\beta}}\left(\frac{G_{r'}}{R_{r'}^\theta}\right)^\lambda\right)^\epsilon}, \tag{A.9}$$

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, Eq. A.9 shares the same structure as the gravity equations based on the spatial models of migration considered in Bryan and Morten (2019) and Morten and Oliveira (2018).