# THREE ESSAYS ON GEOGRAPHY AND TRADE

A Ph.D. Dissertation

by
RAMIZ ABDURAHIMOV

Department of
Economics
İhsan Doğramacı Bilkent University
Ankara
August 2024

To my family and friends

# THREE ESSAYS ON GEOGRAPHY AND TRADE

The Graduate School of Economics and Social Sciences
of
İhsan Doğramacı Bilkent University

by

RAMIZ ABDURAHIMOV

In Partial Fulfillment of the Requirements for the Degree of
DOCTOR OF PHILOSOPHY IN ECONOMICS

THE DEPARTMENT OF
ECONOMICS
İHSAN DOĞRAMACI BİLKENT UNIVERSITY
ANKARA

August 2024

THREE ESSAYS IN GEOGRAPHY AND TRADE
by Ramiz Abdurahimov

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Economics.

____________________

Fitnat Banu Demir
Advisor

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Economics.

____________________

Şaziye Pelin Akyol
Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Economics.

____________________

Erol Taymaz
Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Economics.

____________________

Seda Köymen Özer
Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Economics.

____________________

Başak Tanyeri-Günsür
Examining Committee Member

Approval of the Graduate School of Economics and Social Sciences

____________________

Refet S. Gürkaynak
Director

# ABSTRACT

## THREE ESSAYS ON GEOGRAPHY AND TRADE

Abdurahimov, Ramiz

Ph.D., Department of Economics

Supervisor: Asst. Prof. Dr. Banu Demir

August 2024

Economic activity is unevenly distributed across physical space, influenced by the location, climate, and natural resources of a region. This dissertation contributes to the field of economic geography and quantitative spatial modeling through three essays that address key gaps in the literature.

The first chapter examines the impact of local markets for non-tradable inputs on urban land prices. By using disaggregated spatial data and a two-stage estimation approach with market access as an instrumental variable, I find that districts with 1% more housing used by professional businesses have land prices that are 0.59% higher on average. This estimate increases to 1.24% when adjusted for potential endogeneity, and with robustness confirmed through alternative market access definitions.

The second chapter estimates trade elasticity for Turkish exports at the product level. Utilizing detailed data on import tariffs and trade flows, I analyze the effects of tariff variations on trade flows. The median trade elasticity estimate is -4.48, with differences observed between homogeneous and differentiated goods, as well as between consumption, capital, and intermediate goods. The calculated Turkish welfare gains from trade are 19.8% using a homogeneous aggregate trade elasticity, increasing to 23.8% with heterogeneous trade elasticities in a multi-sector trade model.

The third chapter investigates the effects of Turkey's large infrastructure development program (2005-2015) on the size dispersion within local industries. Improved infrastructure, measured by travel time savings, decreases size dispersion by enhancing access to potential suppliers. This effect reduces trade costs and allows smaller firms to access cheaper inputs. No evidence of pro-competitive pressure affecting size dispersion was found.

Overall, this dissertation advances our understanding of urban economic outcomes, international trade, and domestic trade by addressing significant gaps in economic geography and spatial modeling literature.

*Keywords:* domestic Trade, geography, international trade, quantitative spatial models, urban economics

# ÖZET

## COĞRAFYA VE TİCARET ÜZERİNE ÜÇ MAKALE



Abdurahimov, Ramiz

Doktora, İktisat Bölümü

Tez Yöneticisi: Yrd. Doç. Dr. Fitnat Banu Demir



Ağustos 2024



Ekonomik faaliyetler, bir bölgenin konumu, iklimi ve doğal kaynakları tarafından etkilenerek fiziksel alana düzensiz bir şekilde dağılmıştır. Bu tez, ekonomik coğrafya ve nicel mekansal modelleme alanına, literatürdeki önemli boşlukları ele alan üç makale ile katkıda bulunmaktadır.

Birinci bölüm, yerel piyasaların takas edilemeyen girdiler için kentsel arsa fiyatları üzerindeki etkisini incelemektedir. Ayrıştırılmış mekansal veriler ve piyasa erişimini bir araç değişken olarak kullanarak iki aşamalı bir tahmin yaklaşımı kullanarak, profesyonel işletmeler tarafından kullanılan konutların %1 daha fazla olduğu ilçelerde arsa fiyatlarının ortalama %0.59 daha yüksek olduğunu buluyorum. Bu tahmin, potansiyel içsellik için ayarlandığında %1.24'e yükselmekte ve alternatif piyasa erişim tanımları ile sağlamlığı doğrulanmaktadır.

İkinci bölüm, Türk ihracatı için ürün bazında ticaret esnekliğini tahmin etmektedir. İthalat tarifeleri ve ticaret akışlarına ilişkin ayrıntılı veriler kullanılarak, tarife varyasyonlarının ticaret akışları üzerindeki etkileri analiz edilmektedir. Ortanca ticaret esnekliği tahmini -4.48 olup, homojen ve farklılaştırılmış mallar ile tüketim, sermaye ve ara malları arasında farklılıklar gözlemlenmektedir. Homojen bir toplam ticaret esnekliği kullanarak hesaplanan Türk ticaretinden elde edilen refah kazançları %19.8 olup, çok sektörlü bir ticaret modelinde heterojen ticaret esneklikleri ile %23.8'e yükselmektedir.

Üçüncü bölüm, Türkiye'nin büyük altyapı geliştirme programının (2005-2015) yerel endüstriler içindeki boyut dağılımı üzerindeki etkilerini araştırmaktadır. Seyahat süresi tasarrufları ile ölçülen iyileştirilmiş altyapı, potansiyel tedarikçilere erişimi artırarak boyut dağılımını azaltmaktadır. Bu etki, ticaret maliyetlerini azaltmakta ve daha küçük firmaların daha ucuz girdilere erişmesini sağlamaktadır. Boyut dağılımını etkileyen rekabet baskısına dair bir kanıt bulunmamıştır.

Bu tez, ekonomik coğrafya ve mekansal modelleme literatüründeki önemli boşlukları ele alarak kentsel ekonomik sonuçlar, uluslararası ticaret ve iç ticaret konularında anlayışımızı ilerletmektedir.

*Anahtar Kelimeler:* coğrafya, kentsel ekonomi, nicel mekansal modeller, uluslararası ticaret, ülkeiçi ticaret

## ACKNOWLEDGEMENTS

Completing this PhD thesis has been a challenging yet incredibly rewarding journey, and I would not have been able to reach this milestone without the support, guidance, and encouragement of many individuals and institutions.

First and foremost, I would like to express my deepest gratitude to my supervisor, Banu Demir. Your unwavering support, endless patience, and constructive criticism have been invaluable throughout my research journey. Your expertise and passion for the subject matter have inspired me to push the boundaries of my knowledge and understanding. I am deeply grateful for the time and effort you invested in helping me shape this thesis.

I would like to extend my sincere thanks to my thesis committee members, Şaziye Pelin Akyol and Erol Taymaz, for their invaluable feedback and suggestions. Your thoughtful insights and encouragement have greatly enriched my work. I would also like to thank Başak Tanyeri-Günsür and Seda Köymen Özer for their valuable comments.

A special thanks goes to my colleagues and fellow PhD candidates at the Department of Economics. The discussions, collaborations, and shared experiences have made this journey more enjoyable and intellectually stimulating. I am particularly grateful to Sıla, Zeynep and Merve, whose camaraderie and support have been a constant source of motivation.

On a personal note, I would like to thank my family and friends for their unwavering support and encouragement. To my parents, Azad and Vafa, my dear brother, Fikret, thank you for your love, patience, and belief in me, even when the road seemed long and uncertain. To my friends, Umut, Göksu, Emir, Cankat, Ömer and Kemal, thank you for being my source of strength, laughter, and sanity throughout this process.

This thesis is dedicated to all of you. Thank you.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# CHAPTER 1

# INTRODUCTION

Economic activity is unevenly distributed across physical space. From the earliest civilizations that flourished along riverbanks to the contemporary global economy, the physical characteristics of a region—such as its location, climate, and natural resources—have played a pivotal role in shaping societies and economic opportunities. The field of economic geography has produced work that studies the interactions between economic agents across geographic space. Since the earliest attempts, works, such as Alonso (1964), Mills (1967), Muth (1971), and later by Fujita & Ogawa (1982), Fujita & Thisse (1996) have established a theoretical framework of agglomeration economies. Starting with the seminal works of Krugman (1992, 1996, 1997), economics geography has played an important role in trade literature. The advent of the boom in the available data, has placed geographical disparities into the center of a wide range of economic research. Seminal works like Glaeser et al. (1992), Redding & Sturm (2008), Combes et al. (2018) focus on the role of geography in shaping urban development, while Combes et al. (2012), Donaldson & Hornbeck (2016), Coşar & Demir (2016) study the effect of interconnectedness of regions on key economic factors, such as total factor productivity (TFP), agricultural production and trade patterns, respectively. This dissertation contributes to various gaps in the literature of quantitative spatial modelling in three essays.

The first chapter aims to provide insights into the impact of local markets of non-tradable inputs on urban land prices. Using disaggregated spatial data, I estimate the office propensity (measured by the number of housing units with operating business licenses per capita) elasticity of land prices. A common empirical issue in these setups is that local housing demand and urban land prices are simultaneously influenced by numerous unobserved factors. To identify the causal relationship, I use a two-stage estimation approach, leveraging the exogenous variation in market access across Turkish urban districts as an instrumental variable (IV). Market access is an aggregate measure of how costly it is for a district to trade with other districts, calculated as the sum of the population of all other districts, weighted by the inverse trade cost with the original district. An improvement in a district's market access would imply that it becomes cheaper to ship products to other districts, especially, districts with a larger population having larger weights. Market access, which measures the cost for a district to trade with other districts, serves as a valid IV because it affects land prices indirectly through its impact on local tradable sector activity.

I find that, among two districts with the same population, the district with 1% more housing used by professional businesses has average land prices estimated to be 0.59% higher. This elasticity estimate increases to 1.24% when adjusted for endogeneity. I confirm the robustness of my results to misspecification by using alternative definitions of market access.

The second chapter I estimate the trade elasticity for Turkish exports at the product level. By utilizing detailed data on import tariffs and trade flows, I determine product-specific trade elasticities and analyze how variations in tariffs across destination countries affect trade flows. This allows us to explore key characteristics of Turkish products and calculate the gains from trade within a diverse industry setting.

My baseline median estimate is -4.48 (PPML), which is comparable with the aggregate trade elasticity estimates in the existing literature. Examining the empirical distribution, I find the median elasticity for homogeneous goods to be slightly smaller than differentiated goods (in absolute terms), and the median elasticity for consumption goods to be slightly smaller than capital and intermediate goods (in absolute terms). I find no considerable distributional differences between durable and non-durable goods. Moreover, products with large shares in total exports, smaller median shipment size and smaller seasonal demand variation respond less to changes in tariffs. Finally, using the baseline median estimate as the homogeneous aggregate trade elasticity, I calculate the Turkish welfare gains from trade (*à la* **?**) to be 19.8%. However, utilizing the access to estimated heterogeneous trade elasticities, these gains become 23.8% when used in a multi-sector trade model.

Finally, in the third chapter, I exploit the large infrastructure development program Turkey has undergone throughout 2005-2015, to investigate the effects of the increased connectivity of Turkish provinces on the size dispersion of firms within local industries. The effects of improved infrastructure, measured in travel time savings, reduce trade costs as it becomes faster/cheaper to ship products across locations. This benefits small firms as they are able to sell their products in more remote markets, while also sourcing inputs from distant locations. Consequently, reduction in trade costs increases competition in both output and input markets, reducing the gap in profitability between large and small firms. I construct variables that allow me to disentangle the effect of the exposure to head-to-head competition in the output markets (Edmond et al. (2015), Asturias et al. (2018), Wu et al. (2023)) and the exposure to potential supplier industries ((Bas & Strauss-Kahn (2015), Liu & Qiu (2016), Feng et al. (2016))).

I find that size dispersion, decreases with more access to potential suppliers. A likely mechanism in play here is that, less connected markets induce high trade costs, that only large firms can bear. Hence, decreases in costs of trading across remote markets allow smaller firms have access to cheaper inputs, reducing the size dispersion. I find no evidence of the pro-competitive pressure from increased connectivity affecting size dispersion.

In summary, this dissertation contributes to filling the gaps in various strands of literature, including outcomes concerning urban economic literature, literature on international trade and domestic trade.

# CHAPTER 2

# OFFICE SPACES AND URBAN COSTS

## 2.1. Introduction

Land is a naturally available physical resource whose supply is inherently fixed. It harbors various degrees of economic activity and plays a significant role in economic development. "Law of one price" dictates that in a frictionless economy, identical goods sold in different locations must sell for the same price. However, data universally shows that this is not the case. Financial frictions, trade frictions, and a long list of other economic factors create considerable heterogeneity in the spatial distribution of land prices.

This study aims to provide insights into the impact of local markets of non-tradable inputs on urban land prices. Using disaggregated spatial data, I estimate the office propensity[1] elasticity of land prices. A common empirical concern in these types of setups is that local housing demand and urban land prices are simultaneously co-determined by numerous unobserved factors. I identify the causal relationship by employing a two-stage estimation approach, exploiting the exogenous variation in cross-section of Turkish urban districts' market access as an

[1]measured by the number of housing units with operating business licenses per capita

instrumental variable (IV). Market access is a measure of how costly it is for a district to trade with other districts and can be considered a valid IV because it affects land prices indirectly, through its effect on the local tradable sector activity. Given two districts with the same population, in the one with 1% more housing used by professional businesses, the average land prices are estimated to be 0.59% higher. This elasticity estimate becomes 1.24% when corrected for the endogeneity. I verify the robustness of my results to misspecification using alternative definitions of market access.

Concentration of economic forces in close geographical locations generates benefits in the form of thick labor markets, markets for non-tradable intermediate goods and services, and productivity and knowledge spillovers (Marshall (1890), Krugman (1991), Combes et al. (2012)). However, these advantages are offset by congestion costs, such as high price of local housing (Combes et al. (2018), Behrens et al. (2017)). Cities are predominantly viewed as the outcome of a trade-off between agglomeration economies and urban costs. Fujita & Ogawa (1982) and Fujita & Thisse (1996) have established this phenomenon as the "fundamental trade-off of spatial economics". In their theoretical framework, Allen & Arkolakis (2014) allow productivity and utility from local amenities to be affected by population density, the extent of which are measured by two key parameters. They conceptualize congestion as "... the disutility of higher housing prices" mainly due to the inelastic supply of land and housing. This study aims to provide a plausible empirical estimate of the said congestion effect.

The growth of the industrial production of tradable goods in an area increases the presence of local non-tradable services. These intermediate inputs are less tradable, making it impossible or very costly to trade across long distances. Such inputs include professional services like legal, financial, medical support, venture capital, and specialized technical/repair/IT support. Moretti (2010) finds that each additional job in manufacturing, generates 1.6 jobs in the local non-tradable sector. Black et al. (2005) find that each additional mining job creates 0.17 non-tradable jobs, while the loss of a mining job costs the non-tradable sector 0.34 jobs.

The main contribution of this study is that this is the first attempt (to the best of my knowledge) to empirically estimate the extent of congestion caused by non-tradable inputs. Understanding this channel is important in justifying and evaluating the performance of numerous policies. Place-based policies target underdeveloped regions to reduce inter-regional income disparities and improve economic efficiency by creating economic incentives to increase industrial production and human capital.[2] However, these programs may have adverse general equilibrium effects, such as increasing land prices due to higher housing demand by professional businesses, partly eroding the gains from these policies.

This study is similar in spirit to Combes et al. (2018).[3] They study the elasticity of land prices with respect to city population, where they treat housing as a special form of consumption commodity. The findings of this study complement theirs as I try to explain the role of housing demand by professional businesses in explaining the spatial variation in urban land prices, where housing, as a commodity, is indirectly a part of the production process.[4]

The rest of the chapter is organized as follows. Section 2.2 describes the data.

[2]Moretti (2011) reviews the related literature in great detail.

[3]Combes et al. (2018) construct a model where consumers derive utility from local amenities at the expense of travel costs. At the spatial equilibrium, consumers' optimal choice of location to reside is directly related to the population elasticity of housing prices.

[4]The concept of market access is central to numerous seminal studies, such as Donaldson & Hornbeck (2016), Hornbeck & Rotemberg (2019), where they focus on how US counties' proximity to major trade routes (railroad networks, water transportation) and markets for tradable commodities (agricultural output and manufacturing output, respectively) explain the spatial variation in land prices, using market access as the primary determining factor. Unlike these studies, I conceptualize the presence of professional businesses as the main driving force shaping urban land prices while borrowing their definition of market access to construct an IV.

Section 2.3 discusses the empirical setup. Section 2.4 reports the estimation results. Section 2.5 concludes with policy discussion.

## 2.2. Data

Türkiye is administratively comprised of 81 provinces that are divided into 973 districts, which consist of 32254 neighborhoods. This study is based on the cross-sectional data from December 2019 and comes from various sources.

*Office Propensity.* The Address Based Population Registry System (ADNKS) by the Turkish Statistical Institute (TurkStat) provides us with data on the number of housing units, and housing units with business licenses at the neighborhood level. These units do not include large industrial constructions, such as factories, warehouses, supermarkets, but are rather houses/apartments that serve as office spaces for mainly professional businesses. Professional businesses in this study refer to companies or organizations that provide specialized services in a particular field, such as law, accounting, consulting, IT, engineering, medicine, etc. The number of housing units with business licenses (normalized by the districts' population) serves as a measure of local demand for housing by professional services.

*Land Prices.* Turkish Revenue Administration reports neighborhood-level land price data. These reported land prices are not transaction prices. Instead, they are calculated as an average based on per-meter-square transaction prices of comparable land lots and can be aggregated as an average at the district level. This measure of land prices serves as a close approximation of transaction prices.[5] Having access to the 2019 and 2014 cycles of this data, I use the former as the

[5] It is important to clarify that I observe land prices in the data but not prices of housing units. When dealing with urban housing, it is difficult to account for unobserved qualitative characteristics of apartments and buildings, whereas land is a far more homogeneous commodity than housing. Therefore, in this study, symptoms of congestion are studied in the form of increased land prices.

dependent variable in my analysis, and the latter for weighting observations to adjust for the heterogeneity of the error term.

*Socio-Economic Features.* The Address Based Population Registry System (ADNKS) by the Turkish Statistical Institute (TurkStat) provides detailed district-level urban demographic statistics such as total population, population by 5-year age groups, and population by education groups.[6] I also use a socio-economic development index (SEDI) released by the Turkish State Planning Organization to control for the heterogeneity in the size of economic activities and social development across districts. The index is calculated as a weighted average of 52 socio-economic indicators such as household income, employment, number of schools, hospitals, etc, within districts.

*Geographical Features.* I also use extensive data on the geographical features of districts across Türkiye, namely the area size of the district and bilateral distance among districts. Data on the shortest distances between districts comes from the official digitized road network maps published by the General Directorate of Highways for 2015. [7]

[6] I have 2014 data on district-level population by 5-year age groups, and population by education groups, which I use in order to sidestep endogeneity problems in these variables.

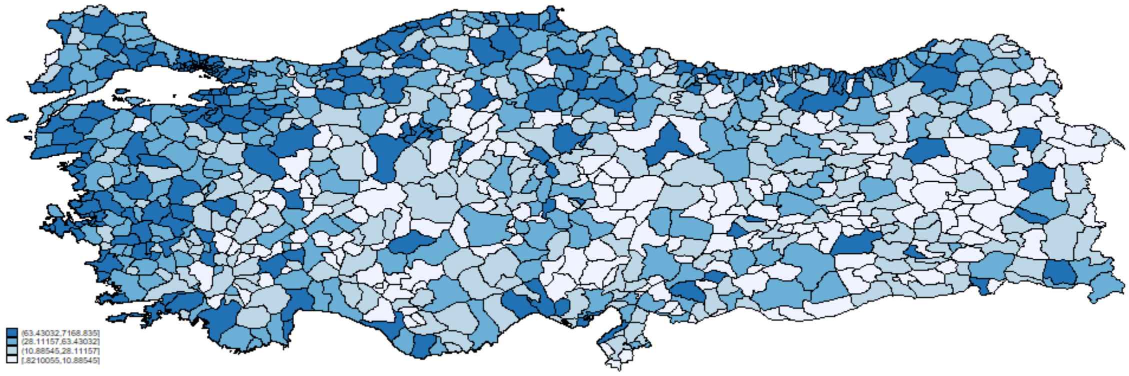
[7] See Coşar et al. (2021*a*) for detailed description on calculations of shortest distances among Turkish districts.

Table 2.1: Descriptive statistics for district characteristics

| Variable | Mean | Std. Deviation | Obs. |
|---|---|---|---|
| Land price ($m^2$) | 357.97 | 837.17 | 953 |
| Population | 79391.72 | 137037.2 | 953 |
| Offices per capita | 0.07 | 0.04 | 953 |
| Offices | 5562.95 | 10692.25 | 953 |
| Housing | 22056.53 | 39106.08 | 953 |
| District area ($km^2$) | 597.92 | 704.89 | 953 |
| Distance to the closest port | 107.62 | 109.73 | 953 |
| Distance to the closes border | 349.74 | 245.33 | 953 |
| SEDI index | 4.10 | 4.23 | 953 |
| Share of working age population in 2014 | 0.68 | 0.05 | 953 |
| Share of college degree population in 2014 | 0.10 | 0.06 | 953 |

The data is complete for 953 (out of 973) districts. Table 2.1 reports the summary statistics for crucial variables at the district level. Based on the spatial distribution of district-level average land prices and offices per capita, illustrated in figures 2.1a and 2.1b, we see similar patterns in both variables, i.e. districts in the western part of the country have high land prices and high density of offices per population.

(a) Distribution of Land Prices across Turkish Districts




(b) Distribution of Office Propensity across Turkish Districts

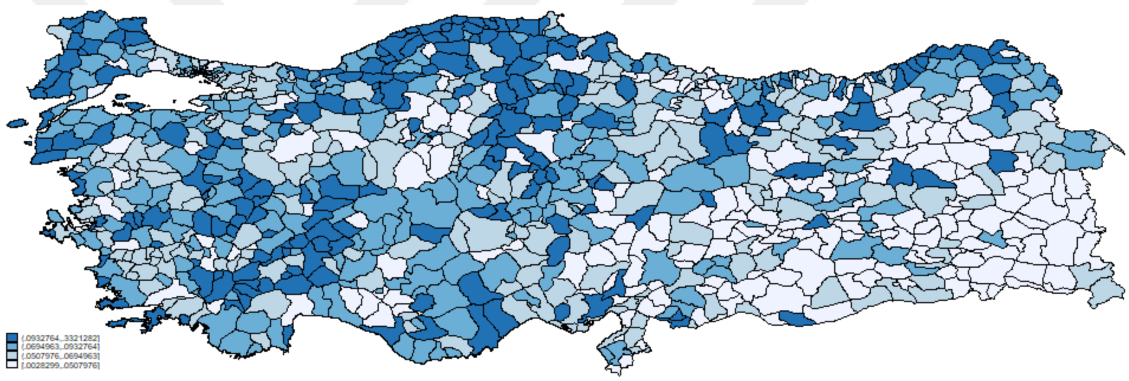



Figure 2.1: Distribution of Land Prices and Office Propensity across Turkish Districts

## 2.3. Empirical Methodology

### 2.3.1. Main Empirical Specification

My main empirical specification is given by:

$$log\,(\text{Land Prices})_i = \beta\,log(\text{Office Propensity})_i + \gamma X_i + \alpha_c + u_i \tag{2.1}$$

where $X_i$ is a vector of controls for each district, and $\alpha_c$ are province-level fixed effects.[8] My focus throughout this study is correctly estimating $\beta$, the office propensity elasticity of urban land prices. The office propensity (*i.e.,* number of offices per capita) is my measure of local non-tradable input demand.[9]

One could potentially estimate equation 4.1 at the province (more aggregate) or neighborhood (less aggregate) level. Neighborhood-level data are remarkably heterogeneous and greatly noisy making statistical inference less reliable.[10]. On the other hand, provinces being large, are subject to different policies based on their economic and socio-demographic significance. This leads to selection bias as not all provinces are administratively treated the same. Therefore, I adhere to using district level data, as it is not as noisy as the neighborhood-level data, and less prone to being specifically targeted by the authorities than provinces.

The office propensity is, conceptually, endogenous as it is simultaneously co-determined with land prices rendering potentially biased least-square estimates. My proposed remedy is to use 2SLS estimation, using market access as an instrumental variable (IV).

#### 2.3.2. Market Access

The literature provides us with a variety of ways of measuring market access. The general form of my measure is:

[8] Istanbul is the largest city in Türkiye, with the population of the next three big cities combined. It lies on both the European and Asian mainland. Therefore, I use separate fixed effects for each half of the city.

[9] Alternatively, we can measure it by the fraction of housing units with business licenses in the pool of all available housing units within a district. The results of the study are robust to the alternative calculation of office propensity (see tables A3 and A4 in the appendix section).

[10] Population of the smallest neighborhood in the dataset is 11 (Manisa, Kobaklar), while the largest neighborhood is 115,000 Manisa, Kobaklar

$$\text{Market Access}_i = \sum_d \tau_{id}^{-\theta} N_d \tag{2.2}$$

where market access of district $i$ is the sum of the population of all other districts ($N_d$), weighted by the penalized inverse trade cost with the original district ($\tau_{id}^{-\theta}$). Trade cost is somewhat of an ambiguous term in the trade literature. The definition ranges from to protectionist regulations to literal freight costs. In this study I use bilateral distance as a proxy for trade costs, $\tau$ This reduced-form expression for market access is theoretically and empirically consistent across many trade models.

Market access is an aggregate measure of how costly it is for a district to trade with other districts. An improvement in a district's market access would imply that it becomes cheaper to ship products to other districts, especially if the other district has a larger population. Meanwhile, land is a non-tradable commodity, so its price is unlikely to be directly affected by market access (*i.e.* market access being uncorrelated with the error term). The population in the origin district, $N_i$, might be endogenously co-determined with land prices in district $i$; that is, there are unobserved district characteristics that simultaneously force land prices higher and attract population. Moreover, I suspect that this endogenous relationship persists in nearby locations. I address this issue by excluding all neighboring districts from the calculation of market access and by including province-level fixed effects and a series of control variables. Also, since I use highly disaggregated spatial data, it is unlikely that individual district pairs have sizeable effects on this aggregate measure, that is, market access. These reasons combined justify that my measure of market access is plausibly exogenous.

Market access is a relevant instrument because it is strongly associated with the demand for professional services (as shown in the first-stage results). Improving market access is likely to lead to an increase in tradable economic activity in the area, which would, in turn, lead to a greater demand for professional services. Market access is expected to affect land prices only through its impact on the use of housing by professional businesses, hence satisfying the exclusion restriction assumption.

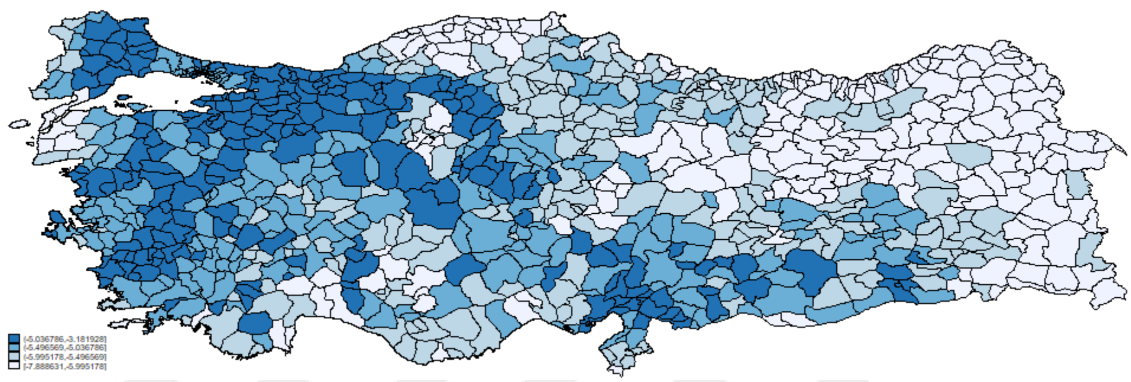



Figure 2.2: Distribution of Market Access across Turkish Districts

Before proceeding further, it is a good practice to visually examine the data. Figure 2.2 plots the spatial distribution of market access across Turkish districts and is quite similar to the map of office propensity. Figure 2.3 plots average land prices against office propensity of districts, without controlling for any observable features of districts. Inspecting the local polynomial fit, we see that the two variables are positively related, hinting towards congestion.[11]

[11]For expository purposes, the measure of office propensity is converted to -per 100,000 inhabitants in figure 2.3; 1st and 99th percentiles have been trimmed.

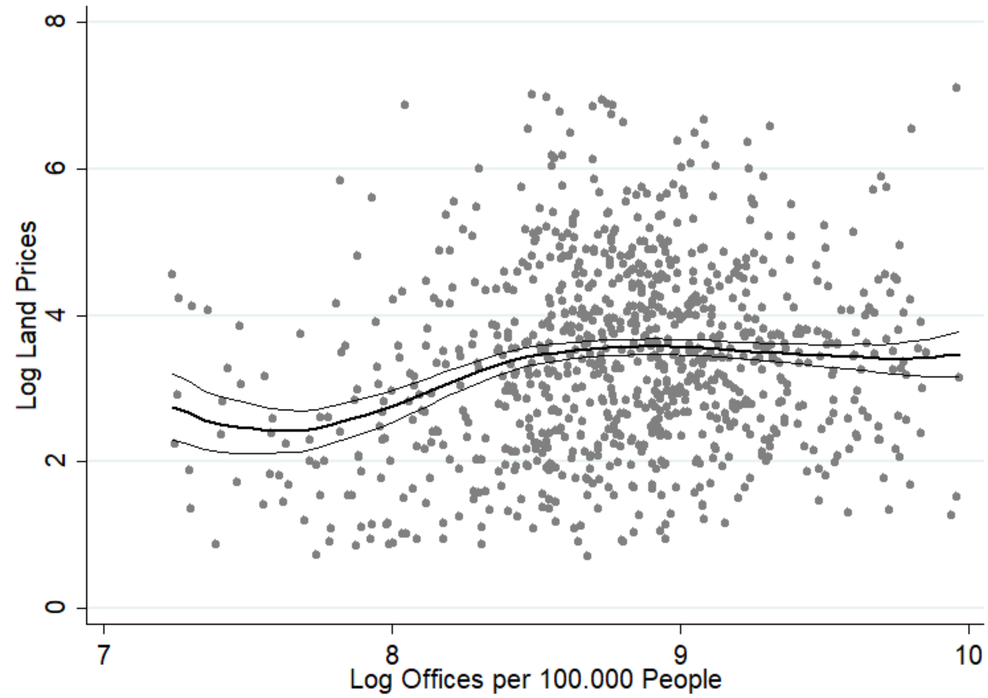



Figure 2.3: Relationship between Land Prices and Office Propensity

## 2.4. Estimation Results

Table 2.2 reports the estimation results of equation 4.1. The vector of control variables used in the regression includes SEDI, share of the working-age population and share of university graduates in the population in 2014, (log) geographical area of the district, (log) district's distance to the urban center and (log) distance to the closest land border and distance to the nearest port.[12] To account for outliers, I follow Donaldson & Hornbeck (2016) and weigh observations in all specifications by each district's average land price in 2014. Standard errors are clustered at the province level to adjust for heteroskedasticity. Column (1) lists the least-squares estimate for $\beta$. Given two districts with the same population, the average land

[12]To preserve the districts whose distances to the closest urban centers, borders, and ports are zero, I use inverse hyperbolic sine transformations as close approximations of the log-distances instead.

prices are expected to be 0.59% higher if one of the districts has 1% more housing used by professional businesses. However, due to the potential endogeneity between land prices and use of housing by professional businesses, the reduced-form estimate is likely to be biased.

Table 2.2: Impact of Office Propensity on Land Prices (Varying Trade Elasticities)

| Dependent Variable: | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Log Land Prices | OLS | 2SLS | 2SLS | 2SLS |
| Log Office Propensity | 0.587** | 1.243*** | 0.855** | 1.323*** |
| | (0.240) | (0.225) | (0.390) | (0.200) |
| First-Stage Coefficient | | 0.441** | 2.105* | 0.247* |
| | | (0.217) | (1.167) | (0.126) |
| First-Stage F-statistic | | 32.31 | 35.05 | 27.74 |
| Observations | 953 | 953 | 953 | 953 |
| Province Fixed Effects | YES | YES | YES | YES |
| Geographical Controls | YES | YES | YES | YES |
| Socio-Economic Controls | YES | YES | YES | YES |

*Notes:* Robust standard errors are clustered at the province level and are reported in parentheses. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively. The detailed list of control variables can be found in the first paragraph of Section 4. Columns (2), (3), and (4) include $\beta$ estimates, where market access in equation 2.2 is calculated with $\theta = 4.14$; $\theta = 1$ and $\theta = 8.22$, respectively, excluding all districts within a 100-km radius.

My definition of market access relies on trade elasticity parameter, $\theta$, and the radius for the exclusion zone. Harris (1954) deems $\theta = 1$ more suitable, while Eaton & Kortum (2002*a*) obtain an average estimate of 8.22. Simonovska & Waugh (2014*a*) develop an alternative methodology and report the average value of $\theta = 4.14$. As a baseline, I choose $\theta = 4.14$ and a 100-km-radius exclusion ring around each district in the calculation of market access.

Column (2) of table 2.2 lists the 2SLS estimate for $\beta$ using baseline market access as the instrumental variable. Given two districts with the same population, in the one with 1% more housing used by professional businesses, the average land prices are expected to be 1.24% higher. These findings are in line with the results of Combes et al. (2018), as they find that a percentage increase in urban population increases land prices by 0.7%.

Columns (3)-(4) of tables 2.2 and 2.3 list the $\beta$ coefficient estimates for equation 4.1, using different values for trade elasticity and alternative radii for exclusion zones, respectively. The baseline specification is robust to alternative choices of both parameters. The first-stage F-statistics are sufficiently large in all specifications, and the relevant first-stage coefficient is statistically significant, while its magnitude varies across specifications. This is due to the mechanical influence alternative values of the parameters have on the interpretation of the first-stage estimates. Despite alternative values of the two parameters having a mechanical effect on the first-stage estimates, the estimated impact of housing use by professional businesses on land prices remains comparable.

Table 2.3: Impact of Office Propensity on Land Prices (Varying Exclusion Zones)

| Dependent Variable: | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Log Land Prices | OLS | 2SLS | 2SLS | 2SLS |
| Log Office Propensity | 0.587** | 1.243*** | 1.034*** | 1.108** |
| | (0.240) | (0.225) | (0.310) | (0.540) |
| First-Stage Coefficient | | 0.441** | 0.239* | 0.306*** |
| | | (0.217) | (0.132) | (0.099) |
| First-Stage F-statistic | | 32.31 | 48.63 | 37.21 |
| Observations | 953 | 953 | 953 | 953 |
| Province Fixed Effects | YES | YES | YES | YES |
| Geographical Controls | YES | YES | YES | YES |
| Socio-Economic Controls | YES | YES | YES | YES |

*Notes:* Robust standard errors are clustered at the province level and are reported in parentheses. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively. The detailed list of control variables can be found in the first paragraph of Section 4. Columns (2), (3), and (4) include $\beta$ estimates, where market access in equation 2.2 is calculated with $\theta = 4.14$, excluding all districts within a 100-km, 75-km, and 125-km radius, respectively.

## 2.5. Conclusion

This study aims to provide insights about the role of non-tradable input markets in shaping urban land prices. Exploiting an exogenous variation in the form of market access, I estimate how much land prices change as a response to an increase in the use of housing by professional business services. I find that a percentage increase in local office propensity increases land prices by 0.59%, controlling for numerous geographical and socio-economic factors. The estimated effect is twice larger (1.24%) when corrected for the endogeneity bias. I verify the robustness of my results to alternative definitions of key variables.

The estimates I report for congestion costs are informative for urban policy-makers. To reduce inter-regional income disparities and improve economic efficiency, many governments implement place-based policies, typically through subsidies and tax incentives. These policies, however, impose large costs on the residents, one of them being high price of land and housing. Although in the long run, the housing supply may (at least partly) adjust to satisfy the increased demand, the local population, in the short run, suffers from higher housing prices. My estimates suggest that these short-run costs are sizeable. An apt urban policy would aim to mitigate these costs via rental subsidies or tax credits.

# CHAPTER 3

# TARIFF-BASED TRADE ELASTICITIES OF TURKISH EXPORTS

## 3.1. Introduction

Trade elasticity is a critical parameter to estimate across many trade models. It is defined as the response of trade flows to changes in trade costs. In a wide range of quantitative trade models, trade elasticity, along with import penetration serve as two main sufficient statistics to calculate gains from trade (Arkolakis et al. 2012). Traditionally trade elasticity is treated as an aggregate parameter shared across products and industries in an economy.

In this study, I estimate the trade elasticity for the Turkish exports at the *product-level*. Utilizing a detailed data on import tariffs and trade flows, I recover the product-specific trade elasticities and analyze the effects of the variation in tariffs across destination countries on trade flows. This helps us explore some key features of Turkish products and calculate the gains from trade in a heterogeneous industry setting.

Despite being methodologically uncomplicated, the literature estimating heterogeneous trade elasticities is fairly sparse. This mainly stems from the (lack of) availability of detailed data on trade costs. Fontagné et al. (2022*a*) collect and form a dataset containing information on tariff rates charged by countries at the level of Harmonized System 6-digit classification (HS6). Using the same dataset I employ a standard structural gravity equation for bilateral trade and estimate product level trade elasticities at the Harmonized System 4-digit classification (HS4) level. The variations in bilateral tariffs across destination countries, HS6 products (within each HS4 class), as well as across years identify my parameter of interest. First, I estimate the gravity equation using logged OLS. Additionally, to address a series of empirical concerns, I employ the approach proposed by Silva & Tenreyro (2006), utilizing Poisson Pseudo Maximum Likelihood (PPML) estimation. I proceed by analyzing the empirical distribution of product-level trade elasticities across different product groups and product characteristics. Finally, I use my estimates in a multi-sector model (Costinot & Rodríguez-Clare 2014) to calculate the overall Turkish gains from trade.

Based on the empirical distributions, my baseline median estimate is -4.48 (PPML), while alternatively the median OLS estimate is -5.44, which are comparable with the aggregate trade elasticity estimates in the existing literature. Upon closer examination of the empirical distribution, I find the median elasticity for homogeneous goods to be slightly smaller than differentiated goods (in absolute terms), and the median elasticity for consumption goods to be slightly smaller than capital and intermediate goods (in absolute terms). I find no considerable distributional differences between durable and non-durable goods. Moreover, products with large shares in total exports, and smaller seasonal demand variation respond less to changes in tariffs (*i.e.* inelastic). Finally, using the baseline median estimate as the homogeneous aggregate trade elasticity (-4.48), I calculate the Turkish welfare gains from trade (as opposed to autarky) to be 19.8%. However, utilizing the access to estimated heterogeneous trade elasticities, these gains become 23.8% when used in a multi-sector trade model.

Trade elasticity can be estimated through various methods: a demand system (Feenstra (1994)) with product-level price data (Simonovska & Waugh (2014*b*); Giri et al. (2021)), a gravity framework (Caliendo & Parro (2015)), or bilateral tariff data (Fontagné et al. (2022*b*)). The requirement for observed trade costs depends on the chosen identification strategy. Estimating a demand system involves analyzing volume and prices at the most detailed classification level of traded products (Feenstra (1994)) without explicitly considering trade policies. Similarly, Simonovska & Waugh (2014*b*) and Giri et al. (2021) use disaggregated price data from 2004 and 1990, respectively. On the other hand, Caliendo & Parro (2015) estimate sectoral trade elasticities by examining cross-sectional variations in trade shares and applied tariffs across 20 sectors and 30 countries. Fontagné et al. (2022*b*) estimate trade elasticity at the HS6 level by utilizing variations in bilateral tariffs. Consequently, the availability of data on trade costs and prices has historically limited previous estimates of trade elasticities at the product level.

Findings in this study also contribute to the literature on calculating gains from trade, contrasting results from single-sector, aggregate trade elasticity models to multi-sector, heterogeneous models (Arkolakis et al. (2012), Costinot & Rodríguez-Clare (2014)). Bolatto & Moramarco (2023) provide an excellent investigation in this matter. They obtain estimates at various levels of aggregation and find that the welfare gains trade openness are large in multi-sector models (such as Arkolakis et al. (2012), Costinot & Rodríguez-Clare (2014), Ossa (2015)) than single-sector model (numerous variations of Eaton & Kortum (2002*b*)).

Methodologically, this study is closest to Fontagné et al. (2022*b*), as it pertains to estimation of product-level, tariff-based trade elasticities using a structural gravity framework. They are able to estimate the parameter estimates at the HS6 level since they are estimating their equation using the full data matrix on trade flows, whereas I estimate it for Turkish exports hence have to conduct the analysis at the HS4 level.[1] Therefore, their interpretation of their results uncover facts about characteristics of product in the light of international trade. The results of this study, on the other hand, shine some light on the peculiarities of Turkish exports and exporters.

The rest of the chapter is organized as follows. Section 3.2 discusses the data. Section 3.3 lays out the empirical methodology. Section 3.4 analyzes the results. Section 3.5 is dedicated to the calculation of gains from trade. Section 3.6 concludes.

## 3.2. Data

To estimate the product-level trade elasticities we need data on two key variables: bilateral trade flows and bilateral applied tariffs. This data has been compiled by Fontagné et al. (2022*a*) and made available upon request by the authors. The full matrix of trade flows of importing and exporting countries are provided BACI (CEPII) database. Trade flows are measured in current U.S. dollars, at the HS6 level, spanning the 2001-2016 period. The advantage of using this dataset is that the empty cells in the trade matrix are filled using the mirror trade flows. Given any year and product class BACI also double checks the import and export trade flows, *i.e.* import and export values are equal. The dataset also reports trade flows net of transportation and insurance costs, in other words, the

[1] One of the challenges during the estimation procedure (even when estimating the gravity equation at the HS4 level) is the lack observations/variation across bilateral tariff rates.

reported values are Free on Board (FoB).

The other key variable of interest is the import tariff rates between countries. Fontagné et al. (2022*b*) provide the tariff data that is compiled from various sources. The raw data provided by International Trade Center (ITC) reports the minimum between the Most Favoured Nation (MFN) custom duty and the preferential duty if it exists (unilateral preference, bilateral trade and/or customs agreements). The authors aggregate the data up to the HS6 product-level and covert charged tariffs into *ad valorem* terms (*e.g.* a tariff rate of zero would correspond to 1 × *ad valorem*).

Finally, to calculate the gains from trade we need the data on final demand shares. I employ the 2016 release of the World Input-Output Table (WIOT) put together by Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R. and de Vries, G. J. (2015). The dataset consists of the full matrix of trade flows between every country-industry pair, as well as expenditure of each country-industry by final use. These final demand items include final consumption expenditure by households and non-profit organisations serving households, final consumption expenditure by government, gross fixed capital formation, and changes in inventories and valuables.

### 3.3. Empirical Methodology

As discussed in the previous section, the final dataset consists of bilateral FoB trade flows, bilateral tariffs at the HS6 level covering the 2001–2016 period. I employ a standard structural gravity equation for bilateral trade for estimating trade elasticity, where for each (of 1055) HS4 product, I estimate:

$$X_{jkt} = exp[\varepsilon_k ln(1 + \tau_{jkt}) + \alpha_{jt}] \times u_{jkt} \tag{3.1}$$

where $X_{jkt}$ are the USD (FoB) value of Turkish exports to destination country $j$ of HS6 product $k$ in year $t$, $\tau_{jkt}$ are tariff rates (import tariff converted into ad-valorem terms) at destination country $j$ for HS6 product $k$ in year $t$. $\alpha_{jt}$ are destination-year fixed effects and $\varepsilon$ denotes the trade elasticity. The variations in bilateral tariffs across destination countries, HS6 products (within each HS4 class), as well as years identify the parameter of interest, $\varepsilon$.

I estimate the structural gravity equation first using OLS taking logs of both sides and weighting each observation by the level of trade flows. Then, to tackle the issue of heteroskedasticity in the error term and the challenge of zero trade flows due to incomplete data, I employ the approach proposed by Silva & Tenreyro (2006), utilizing Poisson Pseudo Maximum Likelihood (PPML) estimation. This non-linear estimation method serves as my primary and favored estimator for equation (3.3).

Gravity equations are central part of trade research. It captures the empirical recurrence of how trade flows between two regions are proportional to the sizes of the two regions and the (inverse of) bilateral distance. One reason for its popularity is the concise and tractable form of the equation. Another reason is that a wide series of quantitative trade models produce equations consistent with the gravity equation, regardless of their theoretical foundations (Anderson (1979), Anderson & Van Wincoop (2003), Krugman et al. (1980), Eaton & Kortum (2002*b*), Melitz (2003)).

Additionally, Arkolakis et al. (2012) show that these models capture the same implications for the relationship between trade flows and the welfare gains can be reduced down to depend on two key economic parameters: responsiveness of trade flows to trade costs (trade elasticity); and trade openness (import penetration).

Consider any trade model that produce the following expression for trade flows:

$$X_{ij} = \frac{F_{ij} w_i^{\varepsilon}}{\sum_{k \in S} F_{kj} w_k^{\varepsilon}} Y_j \tag{3.2}$$

where $X_{ij}$ are trade flows from country $i \in S$ to country $j \in S$, $F_{ij}$ are exogenous country parameters, $\varepsilon < 0$ is trade elasticity, $w_i$ are wages in $i$, $Y_j$ is the GDP in country $j$, and the consumer price index can be written as:

$$P_j = \sum_{k \in S} \left( F_{kj} w_k^{\varepsilon} \right)^{\frac{1}{\varepsilon}} \tag{3.3}$$

Define $\lambda_{ij} \equiv \dfrac{X_{ij}}{Y_j}$, as the share of expenditure of country $j$ on goods exported from country $i$. Then $\lambda_{jj} = 1 - \lambda_{ij}$ is the share of expenditure on domestically produced goods, that serves as a measure of trade openness. From equations 3.3 and 3.3, it follows that for country $j$:

$$\lambda_{jj} = \frac{F_{jj} w_j^{\varepsilon}}{\sum_{k \in S} F_{kj} w_k^{\varepsilon}} = F_{jj} \left( \frac{w_j}{P_j} \right)^{\varepsilon} \tag{3.4}$$

Defining $W_j \equiv \dfrac{w_j}{P_j}$ as the real wages in $j$, one can write:

$$\lambda_{jj} = F_{jj} W_j^{\varepsilon} \implies W_j = \left( \frac{\lambda_{jj}}{F_{jj}} \right)^{\frac{1}{\varepsilon}} \tag{3.5}$$

which implies that the welfare country $j$ depends on trade openness ($\lambda_{jj}$), Trade elasticity ($\varepsilon$) and exogenous domestic parameters ($F_{jj}$). So the welfare impact of any foreign shock to the model (that does not affect $F_{jj}$) can be written as:

$$\hat{W}_j = \left(\hat{\lambda_{jj}}\right)^{\frac{1}{\varepsilon}} \tag{3.6}$$

Arkolakis et al. (2012) also find welfare equations similar to 3.3 across different trade models (monopolistic competition, multi-sector, etc.). In section 3.5, I calculate the welfare gains for the Turkish economy, contrasting the results out of a single-sector vs. a multiple-sector model.

## 3.4. Estimation Results

### 3.4.1. Baseline Results

This section presents trade elasticity estimates for the 1215 HS4 products. The main challenge in this estimation procedure is the limited variation in tariffs, considering narrowly-defined product groups and lack of trade flows. Table 3.1 provides descriptive statistics for estimated trade elasticities. I include $\varepsilon_k$ for HS4 products for which the estimated parameter is strictly negative (guided by the theory) and are statistically significant at the 5% level. After trimming the elasticity estimates that are unsound theoretically and statistically, we are left with 702 HS4 products. Figure 3.1 displays the empirical distribution of trade elasticity parameters, estimated using PPML. The distribution is skewed toward zero, the average $\varepsilon$ being -6.87 and median being -4.48. These distributional moments are numerically comparable to those of Fontagné et al. (2022*b*), which they find the average to be -7.6 and median to be -5.6. Figure 3.2 depicts the correlation between estimated Turkish elasticities and global elasticities reported in Fontagné et al. (2022*b*). Global HS4 elasticities are calculated as average HS6 elasticity within each HS4 heading.

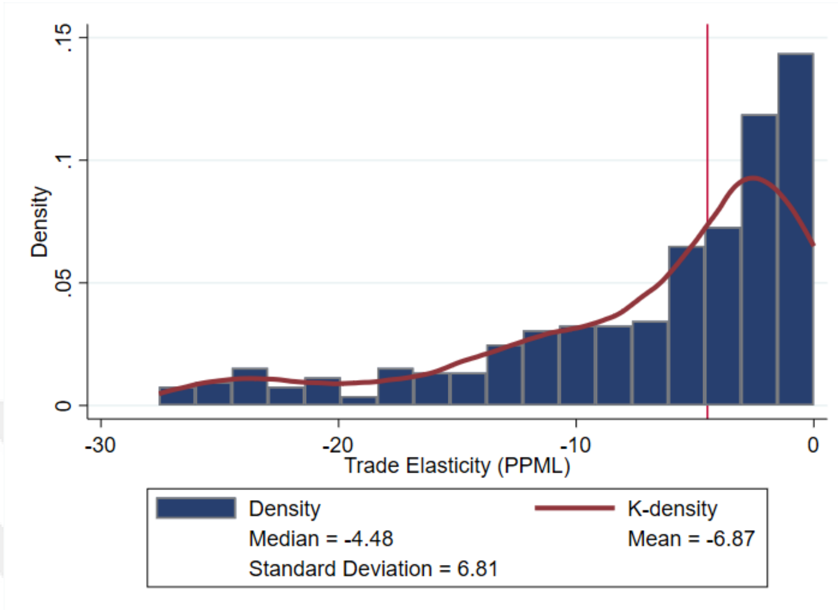



Figure 3.1: The empirical distribution of trade elasticities across HS 4-digit headings, obtained via PPML.

Notes: The graph has been restricted to the sample of products with elasticity estimates not exceeding 30, for expository purposes.

Next, Figure 3.3 depicts the empirical distribution of HS4-digit elasticities obtained via OLS. Applying the same trim used at Figure 3.1 we end up with 758 trade elasticity estimates (strictly negative and statistically significant at the 5% level). The median of the estimates is -5.44, while the mean is -7.61. The standard deviations of the empirical distributions of PPML and OLS estimates are similar (6.81 and 6.98, respectively). Figure 3.4 draws the densities for the two empirical distributions in one graph.

Table 3.1: Descriptive statistics for trade elasticity estimates for across products.

| HS2 Section | Description | Average | Median | Std. Dev. | No. of HS4 products |
|---|---|---|---|---|---|
| 1-5 | Animal & Animal Products | -5.36 | -4.80 | 6.10 | 12 |
| 6-15 | Vegetable Products | -9.49 | -3.05 | 16.50 | 50 |
| 16-24 | Foodstuffs | -3.53 | -2.60 | 3.01 | 40 |
| 25-27 | Mineral Products | -38.88 | -9.40 | 74.66 | 16 |
| 28-38 | Chemicals & Allied Industries | -19.91 | -8.65 | 32.51 | 93 |
| 39-40 | Plastics / Rubbers | -16.76 | -3.25 | 31.08 | 30 |
| 41-43 | Raw Hides, Skins, Leather, & Furs | -27.08 | -9.64 | 36.28 | 20 |
| 44-49 | Wood & Wood Products | -11.51 | -8.96 | 9.89 | 32 |
| 50-63 | Textiles | -13.78 | -4.36 | 28.79 | 107 |
| 64-67 | Footwear / Headgear | -10.14 | -5.55 | 11.36 | 19 |
| 68-71 | Stone / Glass | -19.11 | -14.35 | 21.78 | 46 |
| 72-83 | Metals | -23.48 | -8.24 | 36.76 | 77 |
| 84-85 | Machinery / Electrical | -12.30 | -4.98 | 18.02 | 85 |
| 86-89 | Transportation | -39.76 | -11.86 | 66.94 | 26 |
| 90-97 | Miscellaneous | -15.69 | -5.21 | 32.49 | 49 |

*Notes:* Table lists the mean, median, standard deviation and number of HS4 products of trade elasticities estimated in equation 3.3. The statistics cover 702 non-negative and statistically significant (5% level) HS4 product elasticities.

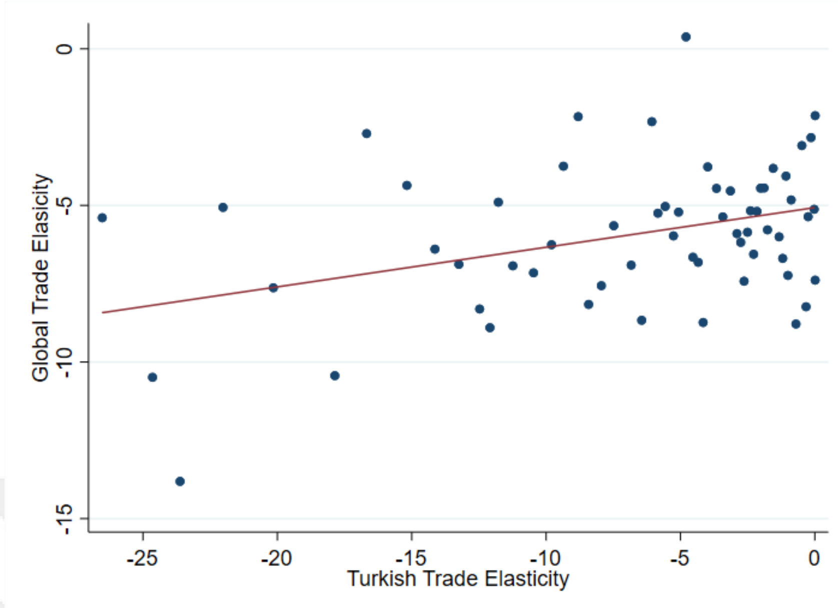



Figure 3.2: Correlation of Turkish and global trade elasticities.

Notes: The horizontal axis represents estimated Turkish trade elasticities. The vertical axis represents depicts elasticity estimates from Fontagné et al. (2022*b*). Global HS4 elasticities are calculated as average HS6 elasticity within each HS4 heading. The graph has been restricted to the sample of products with elasticity estimates not exceeding 30, for expository purposes.

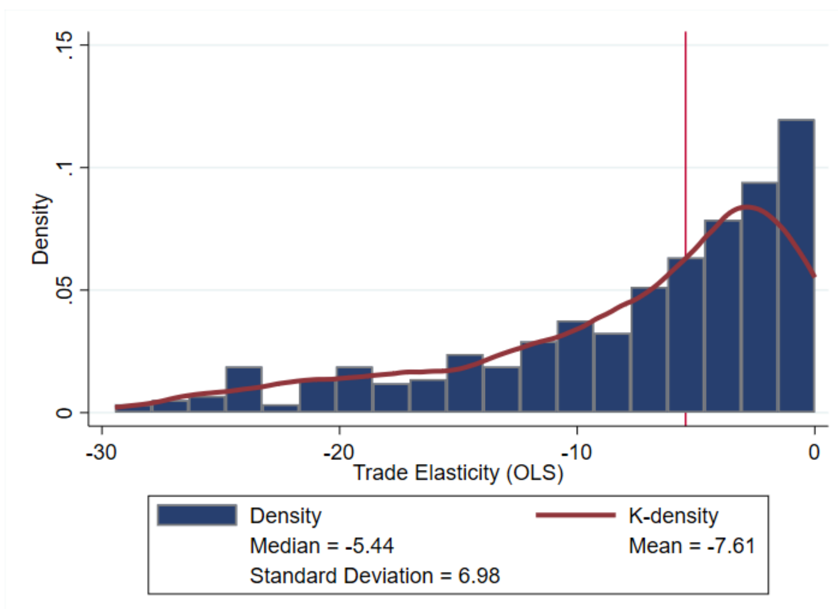



Figure 3.3: The empirical distribution of trade elasticities across HS 4-digit headings, obtained via OLS.

Notes: The graph has been restricted to the sample of products with elasticity estimates not exceeding 30, for expository purposes.

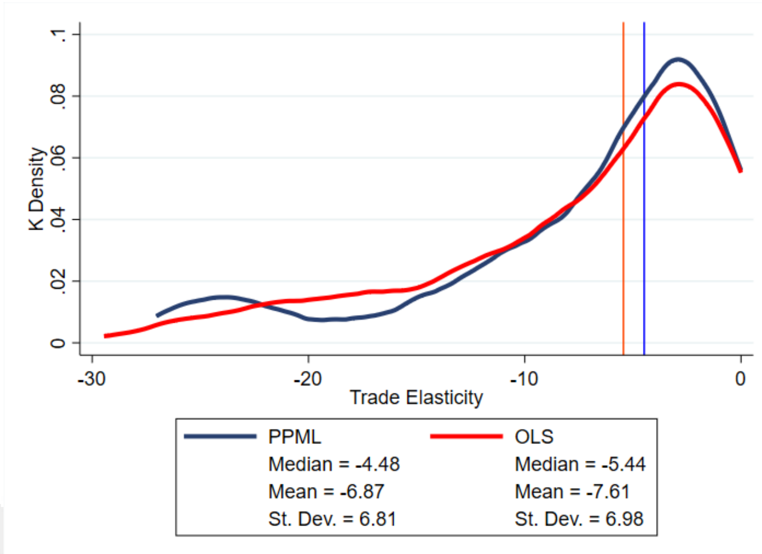



Figure 3.4: The kernel densities of trade elasticities across HS 4-digit heading: PPML vs OLS estimates.

Notes: The graph has been restricted to the sample of products with elasticity estimates not exceeding 30, for expository purposes.

### 3.4.2. Trade Elasticities across Product Groups

Next, I examine how the trade elasticity estimates vary across different classes of products. Moving on, I choose PPML elasticities as my main estimates.

#### 3.4.2.1 Differentiated vs. homogeneous products

To start, figure 3.5 illustrates the kernel density estimates of trade elasticities pertained to homogeneous goods and differentiated goods. I follow the Rauch classification of differentiated vs. homogeneous products. The median elasticity estimate for homogeneous goods is -2.96, while the median elasticity is -4.25 for differentiated goods. A lot of essential raw materials used as inputs of production are homogeneous goods, hence tariffs' effect on trade are more limited than differentiated products. This finding disagrees with the results of Fontagné et al.

(2022*b*) (Table 6). They find that the trade elasticity of differentiated products are smaller (in absolute value), due to the competitive advantage large brands have that dominate their markets. One reason that these two results diverge might be that, studied separately, Turkish exporters fall behind when competing with producers from across the world, where the most-productive firms are able to export to distant markets and face inelastic enough demand. Large exporters from developed countries often have strong, established brands, which are associated with higher quality, or prestige in consumers' eyes, making them less price-sensitive. Exporters from developed countries often have well-established distribution networks, strong relationships with retailers, and better access to marketing channels. These factors can mitigate the impact of tariffs, as these firms have more ways to maintain their market share even if prices rise, whereas a tariff that raises the cost of Turkish exports might push them out of these retail channels.

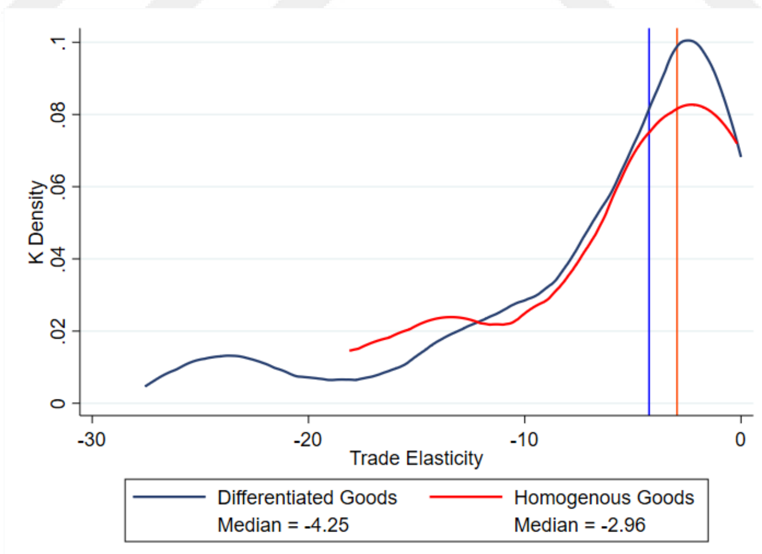



Figure 3.5: The kernel densities of trade elasticities across HS 4-digit headings: Homogeneous vs. Differentiated products.

Notes: The graph has been restricted to the sample of products with elasticity estimates not exceeding 30, for expository purposes.

#### 3.4.2.2 Capital vs. Intermediate vs. Consumption goods

Next, we examine the distributions of trade elasticity across products by use. Figure 3.6 illustrates the kernel densities of capital goods, intermediate goods and consumption goods. The capital and intermediate goods exhibit similar distributions with medians of -4.91 and -6.07. The final consumption goods, however have a steeper distribution with the median trade elasticity of -2.59. It is not surprising, that consumption goods face larger and more inelastic demand at the destination markets, than capital goods and intermediate goods.

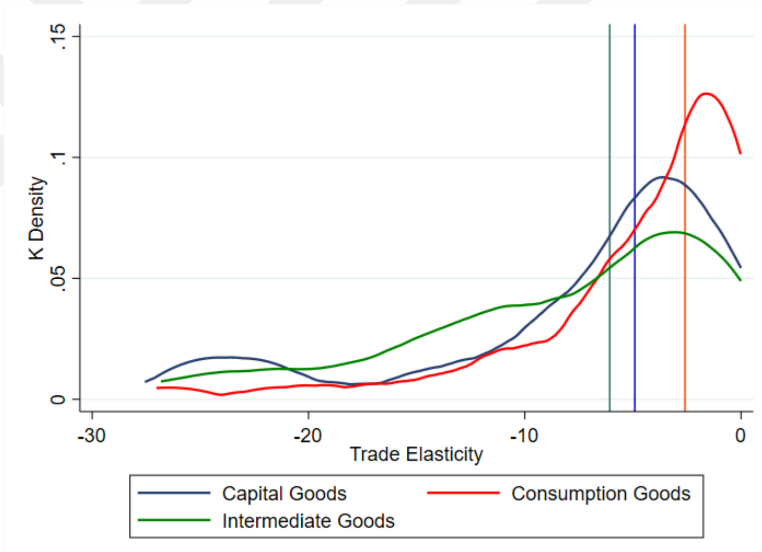



Figure 3.6: The kernel densities of trade elasticities across HS 4-digit heading: Capital goods vs. Intermediate goods vs. Consumption goods.

Notes: The graph has been restricted to the sample of products with elasticity estimates not exceeding 30, for expository purposes.

#### 3.4.2.3 Durable vs. non-durable goods

Now we look at the distribution of trade elasticities across products grouped by durability. Figure 3.7 illustrates the kernel density estimates of durable and non-durable goods. Upon visual examination, the kernel densities look similar with medians being close (durable: -4.91 and non-durable: -4.43).

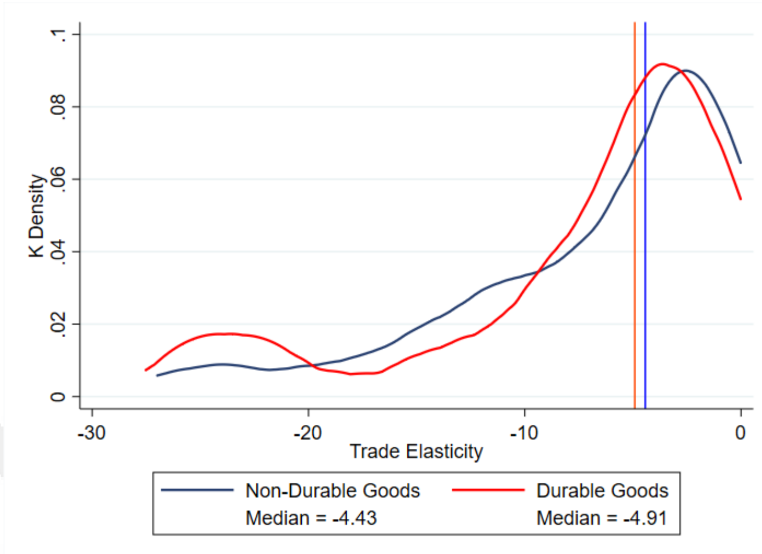



Figure 3.7: The kernel densities of trade elasticities across HS 4-digit heading: Durable vs. non-durable goods.

Notes: The graph has been restricted to the sample of products with elasticity estimates not exceeding 30, for expository purposes.

### 3.4.3. Trade Elasticity and Product Characteristics

In this section, I explore the relationship between a series of product characteristics and trade elasticity. Figure 3.8 illustrates the relationship between the trade elasticity and share of exports within all exports for products under HS4 headings. More inelastic products taking up larger shares of exports is an expected result. Finally, figure 3.9 shows the relationship between the estimated trade elasticity and seasonality of demand across products. Products associated with less seasonality in demand respond less to changes in tariffs.

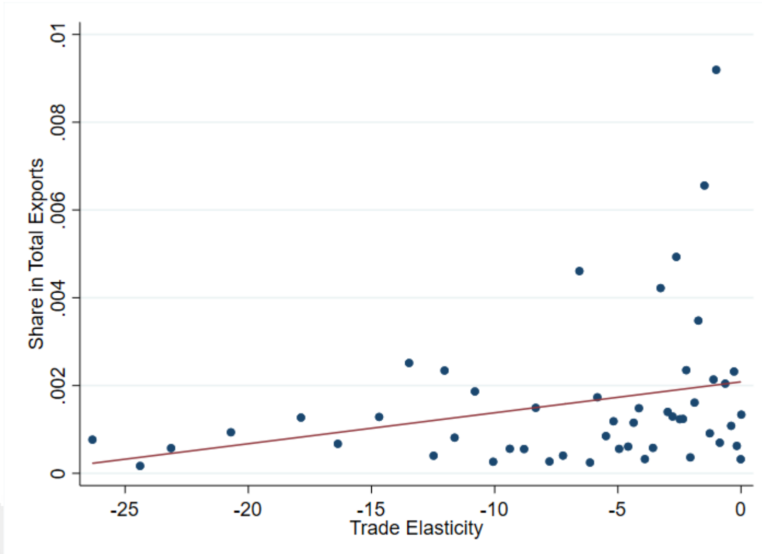



Figure 3.8: Relationship between the share in total exports and trade elasticities.

Notes: Blue dots represent average values within 50 bins of $\varepsilon$. Red line is based on the OLS estimates.

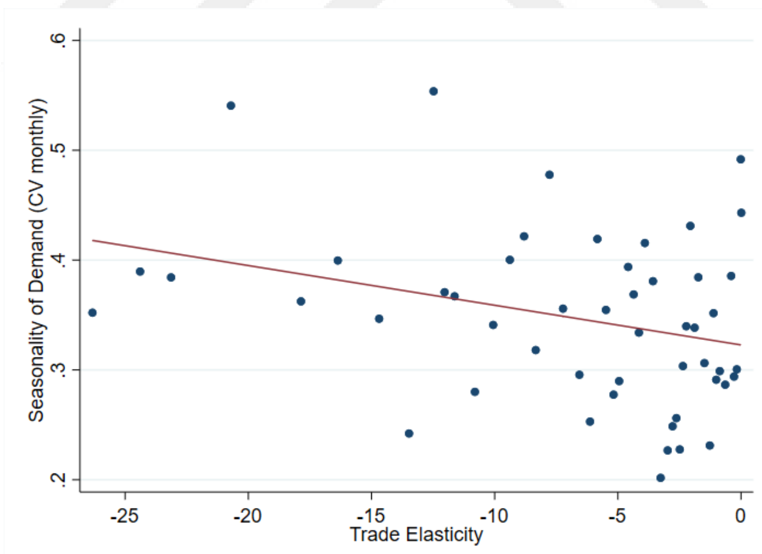



Figure 3.9: Relationship between the seasonality of demand and trade elasticities. Seasonality of demand is calculated as the coefficient of variation across months within a year.

Notes: Blue dots represent average values within 50 bins of $\varepsilon$. Red line is based on the OLS estimates.

### 3.5. Gains from Trade

Finally, I employ Costinot & Rodríguez-Clare (2014) multi-sector model to assess the overall gains from international trade.

$$G = 1 - \prod_{s=1}^{S} \lambda_s^{\frac{\beta_s}{\varepsilon_s}} \tag{3.7}$$

$\lambda_s$ is the share of expenditure on domestic goods, i.e. $\lambda = \dfrac{X_{js}}{X_{js} + X_{is}}$, where $X_{js}$ is the expenditure on domestic goods and $X_{is}$ is the expenditure on imported goods in industry $s$. I define industries according to ISIC 2-digit level. The World Input-Output Table (WIOT) reports trade flows between country-industry pairs. There are 33 large countries (which includes Turkey) and 34th is an aggregate of the Rest of the World.

$\beta_s$ are exogenous preference parameters satisfying $\sum_{s=1}^{S} = 1$. I calculate this term as $\beta_s = \dfrac{F_s}{\sum_{s \in S} F_s}$, where $F_s$ are the total expenditure on final goods produced in industry $s$. On top of 35 ISIC2 manufacturing industries, WIOT also reports total expenditure on final demand at the country-industry level. These include final consumption expenditure by households and non-profit organisations serving households, final consumption expenditure by government, gross fixed capital formation, and changes in inventories and valuables

$\varepsilon_s$ are average trade elasticity of products produced by industry $s$. Using WITS concordance tables to match products (under HS4 classification) to industries that produce them (ISIC2), I calculate the average trade elasticity for each industry. Note that, we are able to calculate trade elasticity only for tradable goods. Following CRC's methodology, for all the industries that produce non-tradables I use $\varepsilon = -4.48$ which is the median trade elasticity across all products.

The gains from trade can also be (and is usually) calculated in a one sector-model where the economy produces one composite good with a homogeneous trade elasticity parameter. In the one sector model the expression in equation 3.7 reduces to:

$$G = \lambda^{\frac{1}{\varepsilon}} \tag{3.8}$$

Table 3.2 reports the calculated gains from trade in 2011 contrasting the results obtained from using a one-sector model and a multiple-sector model.

Table 3.2: Turkish gains from trade

| | (1) | (2) |
|---|---|---|
| | Heterogeneous $\varepsilon$ | Homogeneous $\varepsilon$ |
| CRC Gains from trade | 0.11938 | 0.028798 |
| Our Gains from trade | 0.2381 | 0.1982 |

*Notes:* Column (1) reports the gains from trade in a multi-sector model based on equation 3.7. Column (2) reports the gains from trade in a one-sector model based on equation 3.8, where the aggregate trade elasticity equals median trade elasticity from the according studies.

Column (1) lists the welfare gains from trade with heterogeneous industry trade elasticities, while column (2) displays gains from trade with homogeneous trade elasticity (constant across industries). Both in the replication of Costinot & Rodríguez-Clare (2014) and my calculations, using homogeneous trade elasticity underestimates the gains from trade. The canonical way of interpreting welfare gains from trade is comparing welfare (usually measured in aggregate consumption) under the existing trade regime to autarky. Referring to this interpretation, Türkiye's aggregate consumption is 19.8% higher than what would have been if Türkiye did not trade (according to the one-sector model). Turkish welfare gains are 23.8% according to the results based on multi-sector model. In the one-sector model, we assume that the economy produces one, composite consumption good with trade characteristics embodied in the import penetration and the trade elasticity parameter. Here a foreign shock would have a uniform effect on welfare. On the other hand, multi-sector model allows import penetration and trade elasticity to vary across industries, implying a foreign shock would have differential impact on welfare based on the reliance of trade of these sectors.

## 3.6. Conclusion

In this study, I estimate the trade elasticity for the Turkish exports at the *product-level* (HS4). Utilizing a detailed data on import tariffs and trade flows, I recover the product-specific trade elasticities and analyze the empirical distribution of product-level trade elasticities across different product groups and product characteristics. The baseline median estimate serves as the homogeneous aggregate trade elasticity (-4.48) as I calculate the Turkish welfare gains from trade to be 19.8%. using the product-level estimated heterogeneous trade elasticities, these gains become 23.8% when used in a multi-sector trade model.

# CHAPTER 4

# DO INFRASTRUCTURE UPGRADES AFFECT MARKET DISTORTIONS?

Economists and policymakers widely recognize that investing in transport infrastructure is vital for economic development. The core idea is that efficient transport infrastructure can improve firms' access to external markets and facilitate inter-regional resource mobility, which are essential for regional economic growth and increases in real income (Donaldson & Hornbeck (2016), Hornbeck & Rotemberg (2019), Coşar & Demir (2016)). Enhanced transport infrastructure can lower trade costs, thereby reducing inter-regional price gaps and boosting trade flows. This ultimately leads to a better trading environment, improved market integration, and substantial welfare benefits, such as higher per capita GDP and increased real income.

The dispersion of firm markups within industries is an indicator of possible efficiency gains (see Hsieh & Klenow (2009)). Potential efficiency is attained in perfectly competitive markets, which are associated with uniform markups (zero markups under perfect competition and constant markups under monopolistic competition). Increases in markup variability (henceforth dispersion) reduces the allocative efficiency in the market as large firms with higher markups can get away with charging higher prices. This decline in efficiency is more pronounced in monopolies.

In this study I exploit the large infrastructure development program Turkey has undergone throughout 2005-2015, to investigate the effects of the increased connectivity of Turkish provinces on the firm size dispersion within local industries. The effects of improved infrastructure is measured in travel time savings as a result of the infrastructure development program. I construct variables that allow me to disentangle the effect of exposure to head-to-head competition and exposure to potential supplier industries. I conduct a series of reduced form empirical exercises to analyze the effect of increased exposure to other markets (as a result of the infrastructure upgrades) on size disparities among firms.

To separate the own-industry pro-competitive effect from cross-industry cost advantage effect, I construct three variables that quantify the exposure of firms to supply and market capacities of firms in other provinces. For each province-industry pair these variables take the form of sums of (inverse) market-size-weighted travel times across destination provinces. I employ a gravity equation approach to recover the initial supply and demand capacities of industries, which roughly correspond to export and import capabilities of province industries, respectively. I then regress the (time) difference of the firm size dispersion measure on the (time) difference of these variables, to estimate the impact of each source of gains from improved infrastructure.

My identification strategy is based on the variations in travel times among provinces. Since the program's aim is "ensuring the integrity of the international and national networks and addressing capacity constraints that lead to road traffic accidents", and not focused on trade outcomes, it is plausible to consider the infrastructure investments as exogenous when evaluating the market distortions caused by enhanced connectivity.

I find that firm size dispersion, *i.e.* efficiency losses, decreases with more access to potential suppliers. A likely mechanism in play here is that, less connected markets induce high trade costs, that only large firms can bear. Hence, decreases in costs of trading across remote markets allow smaller firms have access to cheaper inputs, reducing the firm size dispersion. I find no evidence of the pro-competitive pressure from increased connectivity affecting firm size dispersion.

The existing literature has identified two main channels explaining the link between trade liberalization and welfare gains. The productive efficiency channel from the classical Ricardian effect of trade where firms charge prices equal to their marginal cost (Donaldson & Hornbeck (2016), Donaldson (2018)). The allocative efficiency channel is driven by the pro-competitive effects of trade. Specifically, inefficiencies and productivity losses caused by resource misallocation in less competitive markets can be minimized when firms face increased direct competition. An integrated market facilitated by better transport infrastructure allows more competitors to enter, pressuring prices downward. These welfare gains from trade, achieved through the allocative efficiency channel, result from reduced markups and decreased markup dispersion (Edmond et al. (2015), Asturias et al. (2018), Wu et al. (2023)). This study relates to this strand of literature as I am interested in outcomes related to the market distortions. However, the allocative efficiency literature mainly focus on the competitive pressure in the output markets. The current literature has paid little attention to the source of efficiency gains being increased access to imported intermediate inputs (Bas & Strauss-Kahn (2015), Liu & Qiu (2016), Feng et al. (2016)). In my empirical setting, I simultaneously study how increased competitive pressure and easier access to intermediate inputs affect inefficiencies that arise due to market distortions.

Existing literature rely on the dispersion of firm total factor productivity (TFP) and markups to measure the extent of misallocation. In this study, I focus on firm size rather than TFP and markups. Practically, firm size (measured either with the number of employees or the value gross sales) is a good proxy of TFP and markups, as larger firms often have more market power, which allows them to set higher prices relative to their marginal costs, resulting in higher markups. This is because larger firms typically enjoy economies of scale, and lower costs enabling them to sustain higher markups. De Loecker et al. (2020) found that markups have been rising globally, particularly in the United States, and that this trend is correlated with increasing disparities in firm size, contributing to inefficiencies in resource allocation. Aghion et al. (2005) suggest that larger firms with market power can stifle innovation and reduce dynamic efficiency. This, in turn, has negative implications for allocative efficiency, as markets dominated by large firms may not allocate resources towards the most productive and innovative uses.

This analysis could be extended to measure the markup dispersion to directly relate to the literature on markup dispersion. Meanwhile, as firm markups and size are correlated, this study complements the growing literature discussed above.

The rest of the chapter is organized as follows. Section 4.1 provides background on the infrastructure development program of 2000s in Turkey. Section 4.2 discusses the data. Section 4.3 lays out the empirical strategy and construction of key variables. Section 4.4 reports the baseline estimation results and a series of robustness checks. Section 4.5 concludes.

## 4.1. Background

Turkey is administratively divided into 81 provinces. It is a large country with significant internal trade costs, as the average distance between its provinces is nearly 765 km, ranging from 45 km to 2050 km. This results in considerable costs for transporting goods between different locations. Roads are the primary mode of domestic transportation, handling a large portion of domestic freight and passenger traffic. This reliance on road transportation drove authorities to make substantial investments in infrastructure during the 2000s. Even though Turkey already had an extensive road network connecting provincial centers by 2005, most segments lacked dual carriageways, leading to limited capacity and long-considered inadequate infrastructure.

The Turkish government initiated a large-scale transportation investment program in 2002 aimed at ”ensuring the integrity of the national network and addressing capacity constraints that lead to road traffic accidents.” This investment led to a significant conversion of single carriageways (undivided two-lane roads) into dual carriageways. The increased capacity enabled vehicles to travel more reliably at higher speeds, resulting in more predictable arrival times and reduced accident rates, including a decrease in fatalities per kilometer traveled.

Documentary evidence of the investment objectives and design does not suggest that road segments were selected based on domestic trade-related outcomes. Firstly, policy documents emphasize the long-term goal of improving connections between all provincial centers to create a comprehensive grid network across the country, rather than enhancing trade between specific regions. The General Directorate of Highways policy outlines criteria focused on ”ensuring the integrity of the international and national networks and addressing capacity constraints that lead to road traffic accidents." Secondly, the extent of road upgrades varied significantly across provinces, with no visible concentration in specific regions. Lastly, the investment was centrally planned and financed by the central government, with no direct involvement from local administrations.

My identification strategy relies on the variations in connectivity across provinces. Given the non-trade oriented goal of the program, it is plausible to assume the exogenous nature of the investments in infrastructure when assessing the market distortions as a result of improved connectivity.

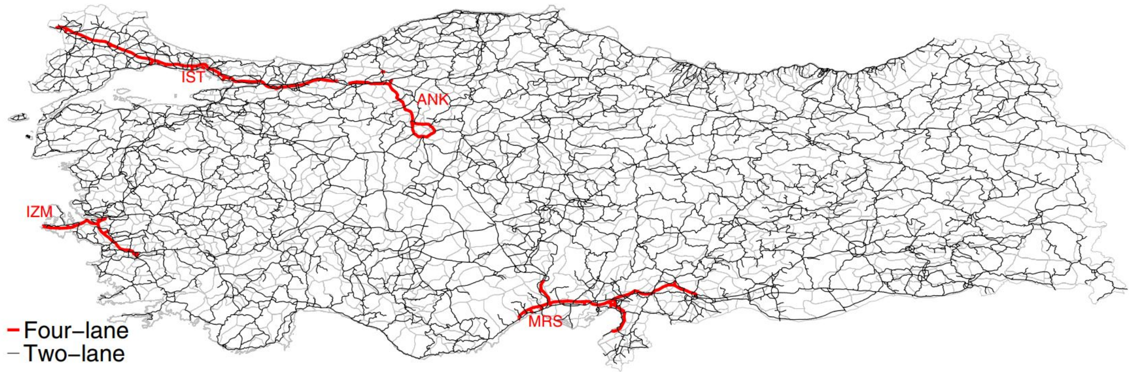
### 4.2. Data

A notable aspect of this study is the access to high-quality data on detailed firm characteristics and domestic trade flows within Turkey during the period of significant road network upgrades. This data, sourced from the Turkish Ministry of Industry and Technology (MoIT), includes administrative firm-to-firm transaction records. Since 2006, Turkish firms have been legally required to report all transactions with a business partner if the annualized amount exceeds a certain threshold to the Ministry of Finance (*e.g.* this amount was (USD 3300 in 2010). This requirement aims to reduce tax evasion and increase VAT collection. Each transaction report is cross-checked to account for the cases of inconsistencies. The dataset also provides information on the number of employees, industry and location of the firm.

The other important part of this study relies on the quantification of the benefits of the road upgrades. I utilize the data used in Coşar & Demir (2016), Coşar et al. (2021*b*), Demir et al. (2024). They calculated the decadal change in inter-provincial travel times. They digitized official road network maps published by the General Directorate of Highways (GDH) for 2005, 2010, and 2015. Using geographic information system (GIS) software, they determined the fastest possible travel times between the 81 provincial centers for each year. On single-lane carriageways, where congestion is common they assume the average travel speed (reported by GDH) of 65 km/h. On expressways and highways where the traffic runs smoothly, they assumed the maximum legal driving speeds of 90 km/h and 110 km/h, respectively. The average travel time between any two provinces decreased by 1.5 hours, down from an average of 6.5 hours in 2005. Time savings increased with distance, reaching up to five hours for cities 1,500 km or more apart.

Finally, to serve as controls, I use a number of variables on industry characteristics using 2005 UK import data. These variables include each industry's average log weight-to-value ratio; time sensitivity, *i.e.* share of fright carried via air transport. Additionally, I calculate British industries' share of expenditure on transportation using World Input Output Table (WIOT) 2011 cycle.

(a) Road Network in 2005




(b) Road Network in 2015

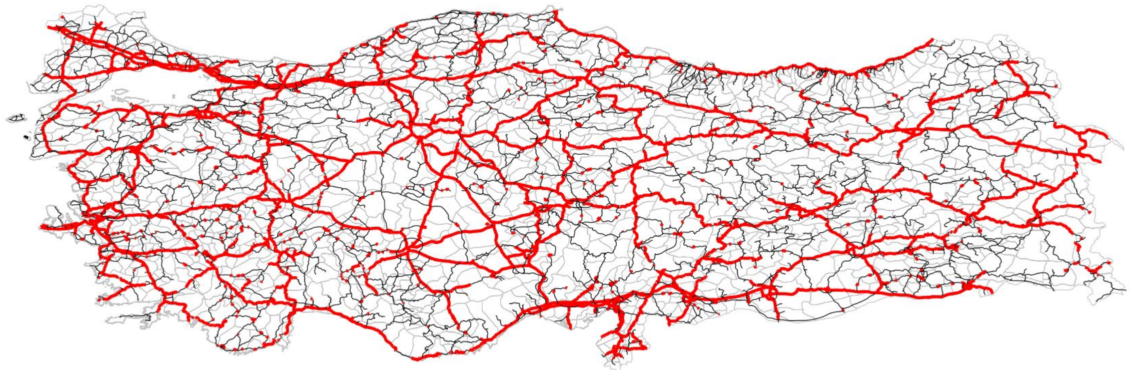


Figure 4.1: Turkish road network.

Coşar et al. (2021*b*), figure 1. Gray lines denote district boundaries. Black lines represent single-carriageway roads and red lines represent dual-carriageway roads (highways and expressways). Four large cities mentioned in the text are labeled in the top panel: IST (Istanbul), IZM (Izmir), MRS (Mersin) and ANK (Ankara).

## 4.3. Econometric Strategy

### 4.3.1. Baseline Specification

The baseline empirical model is given by the following equation:

$$\begin{aligned} \Delta SizeDispersion_{is} = \quad & \beta_1 \Delta Competition_{is}^{output} + \beta_2 \Delta Competition_{is}^{input} \\ & \beta_3 \Delta DemandConnectivity_i + X_i + X_s + \alpha_{is} + \varepsilon_{is} \qquad (4.1) \end{aligned}$$

where time differences in variables are between 2006-2016, $\alpha_{is}$ are province-industry fixed effects, $X_i$ and $X_s$ are province and industry characteristics, respectively. $\varepsilon_{is}$ is the error term. I explain the definitions and measurements of $\Delta SizeDispersion_{is}$, $\Delta Competition_{is}^{output}$, $\Delta Competition_{is}^{input}$ and $\Delta DemandConnectivity_i$ in the following section.

### 4.3.2. Variables

Equations include the subscript $t$ for the ease of notation. For this study $t \in \{2006, 2016\}$ $\Delta$ sign refers to changes in variables across time and will be explicitly described what type of difference is taken (*e.g.* log-differences, percentage differences, etc.). Before proceeding any further, we must make the clarification of timing. Since the firm-level data provided by the Ministry of Industry and Technology starts at 2006, I base my analysis on the 2006-2016 period, whereas the data on road upgrade program spans covers the 2005-2015 period. Considering that these 10-year periods are only one year apart and granted the massive, nationwide scale of the infrastructure program, it is not implausible to equate these periods when discussing causes and effects inferred from this analysis.

#### 4.3.2.1 Firm Size Dispersion

$\Delta SizeDispersion_{is}$ refers to the change in the variation of firm size at the province-industry level across time. I measure the size of firms in two variables: number of employees and sales. There are different methods of measuring how much the size varies within a market. My main measure of firm size dispersion is via coefficient of variation (CV, henceforth):

$$CV_{ist} = \frac{\sqrt{Variance(y)_{ist}}}{\overline{y}_{ist}} \tag{4.2}$$

where $y$ refers to the size of firm in province $i$, industry $s$ at time $t$. An alternative measure of dispersion is based on Theil index, which is given by:

$$Theil_{ist} = \frac{1}{n_{ist}} \sum_{f=1}^{n_{ist}} \frac{y_{fist}}{\overline{y}_{ist}} log\left(\frac{y_{fist}}{\overline{y}_{ist}}\right) \tag{4.3}$$

where $y$ refers to the size of firm $f$ in province $i$, industry $s$ at time $t$. Both of these indices quantify the amount of firm size disparity within a market. Smaller values correspond to low levels of size disparity between larger firms and smaller firms, hence more competitive environment. Likely, large values of both indices indicate highly dispersed markets, *i.e.* dominant producers being much larger than smaller firms in terms of size.

I measure the time difference ($\Delta$) according to the midpoint formula in the baseline specification (Davis & Haltiwanger (1998)):

$$\Delta X = \frac{Z_{2016} - Z_{2006}}{(Z_{2016} + Z_{2006})/2} \tag{4.4}$$

The advantage of using the midpoint formula is that even if one of the measurements of $Z_t$ or $Z_{t-1}$ is zero, the change is always defined and varies between -2 and 2. I also calculate the change in firm size dispersion in log differences (*i.e.* $\Delta Z = logZ_{2016} - logZ_{2006}$) and re-estimate the main specification, challenging the robustness of the results.

#### 4.3.2.2 Connectivity Measures

The benefit firms get from improved connectivity as the result of the road upgrade program is represented in three terms: $\Delta Competition_{is}^{output}$,

$\Delta Competition_{is}^{input}$ and $\Delta DemandConnectivity_i$.

$$Competition_{ist}^{output} = \sum_{j \neq i} \frac{\widehat{SC}_{js}}{TT_{ijt}} \tag{4.5}$$

$$Competition_{ist}^{input} = \sum_{j \neq i} \frac{\sum_{r \neq s} \widehat{SC}_{jr} \cdot W_r}{TT_{ijt}} \tag{4.6}$$

$$DemandConnectivity_{it} = \sum_{j \neq i} \frac{\widehat{MC}_j}{TT_{ijt}} \tag{4.7}$$

$TT_{ijt}$ refers to travel times between provinces $i$ and $j$ at time $t$, while the estimation procedure of the terms $\widehat{SC}$ and $\widehat{MC}$ will be discussed in detail shortly. Firstly, all of these terms are related to travel times between provinces, where improvements stem from reduction in travel times. Having access to the full matrix of travel times between 81 provinces, measured in 2005 (before the road upgrade program) and 2015 (after the road upgrade program), these connectivity terms serve as a tool to quantify the effects of increased exposure to competition in the output markets, input markets and overall demand.

The terms supply capacity, $\widehat{SC}_{is}$ and market capacity, $\widehat{MC}_i$ (indexed differently) roughly represent the market "sizes" of province-industry pairs. $W_r$ is a weighting matrix that is comprised of purchases of industry $s$ from industry $r$, expressed as percentages within all expenses. Ignoring $t$ for a moment, one may observe that the exposure terms in equations (4.5) - (4.7) are sums of (inverse) market-size-weighted travel times across destination provinces $j$ ($i \neq j$ being the origin). We are interested in the supply capacity of province-industry pairs, because we want to quantify the outcomes of improved connectivity, by separating the own-industry pro-competitive effect from cross-industry cost advantage effect. On the other hand, we do not make distinctions of industries when calculating the market capacity of locations because we are interested in the measure of an overall purchasing/demand capacity of provinces. This term may include demand magnitudes of output markets, input markets, service industries, etc altogether.

$Competition_{ist}^{output}$ quantifies the exposure of a firm in province $i$, industry $s$ and at time $t$ to the supply capacity of the industry $s$ located in other provinces. An increase in this term would represent the pro-competitive pressure experienced by a firm as the result of a more integrated output market due to the increased competition with all other producers of that industry that are located elsewhere. $Competition_{ist}^{input}$, meanwhile, quantifies the exposure of a firm in province $i$, industry $s$ and at time $t$ to the supply capacity of all other industries other than $s$ located in other provinces. An improvement in this term would correspond to increased competition in the market of suppliers, thus potentially reducing the input prices. Finally, $DemandConnectivity_{it}$ quantifies the exposure of a firm in province $i$, at time $t$ to increased access to demand in output markets in other provinces, facilitated by connectivity improvements.

#### 4.3.2.3 Supply Capacity and Market Capacity

In order to obtain these estimates I employ a gravity approach (Redding & Venables (2004)). Using the firm-to-firm transactions data, I construct the full matrix of trade flows between 1944 (81 provinces × 24 manufacturing industries) province-industry pairs $is$ in year 2006. Following Silva & Tenreyro (2006), I estimate the following equation using the Poisson Pseudo-Maximum Likelihood (PPML) method:

$$X_{isj} = exp\{\gamma D_{ij} + ex_{is} + im_j + u_{isj}\} \tag{4.8}$$

where $X_{isj}$ are the trade flows between origin province $i$ and destination province $j$ in industry $s$, while $D_{ij}$ are the bilateral closest road distances between origin province $i$ and destination province $j$, $u_{isj}$ is the error term. $ga\hat{m}maD_{ij}$ captures the effect of distance (that proxies the trade costs) on trade flows, while the estimated exporter-specific $\widehat{ex}_{is}$ and importer-specific fixed effects $\widehat{im}_j$ explain the variation in trade flows that are characteristic to exporters and importers. Here, following the methodology of Redding & Venables (2004), I claim that these fixed effects capture the supply and demand capacities of province-industries. Put differently, variation in trade flows that are not explained by trade costs are attributed to varying capacities of provinces to supply and demand goods. Hence, these fixed effect estimates serve as the supply capacity and demand capacity estimates for calculating the connectivity expressions in equations (4.5)-(4.7).

$$\widehat{SC}_{is} = \widehat{ex}_{is} \qquad \widehat{MC}_i = \widehat{im}_j \tag{4.9}$$

Finally, $\Delta Competition_{is}^{output}$, $\Delta Competition_{is}^{input}$ and $\Delta DemandConnectivity_i$ are nothing but logarithmic differences of these terms across time, *i.e.* $\Delta Z = logZ_{2016} - logZ_{2006}$.

### 4.3.3. An Illustrative Example

In order to illustrate how this variables work, consider an example. Suppose a firm located in Aksaray (a small town) operating in the chemical industry. We assume that the trade costs of freight within a province is zero. As $Competition^{output}$ measures the exposure of this firm to head-to-head competition, an increase in this variable for the firm can stem from two sources: either (i) an increase in the supply capacity of producers in the chemical industry (holding all else constant); or (ii) a reduction in travel times among provinces, reducing the cost of shipping output to other locations, hence enabling the firm (and all other firms in chemicals) to sell in more distant markets. Referring to the situation in (ii), as a result of the infrastructure program, chemical products produced by this firm can now be sold in, say Istanbul along with other chemical producers, whereas prior to the program shipping costs would have been too great for the firm to operate in Istanbul chemical market.

Similarly, for the firm located in Aksaray (a small town) operating in the chemical industry, $Competition^{input}$ measures the exposure of this firm to producers in industries other than chemical. An increase in this variable can stem from two sources: (i) an increase in supply capacities of other industries or; (ii) a reduction in travel times among locations, making it less costly to outsource goods produced by other industries located elsewhere. Suppose that the chemical industry is a major importer of goods from the mineral industry. Referring to the hypothetical scenario described in (i), the amount of total production in the mineral industry allows our firm to buy inputs cheaper. On the other hand, reduction in trade costs (referring to scenario ((ii)), would expand the menu of producers in the mineral industry from which our firm can purchase, picking the cheapest supplier.

Finally, $DemandConnectivity$ measures the amount of demand in other locations. An improvement in this variable can stem from: (i) an increase in the demand capacity of other provinces and; (ii) reduction in travel times among provinces. The increased demand in other provinces would incentivize our firm to operate in other markets, while the reduction in travel times would enable our firm to meet the demand in more remote markets.

### 4.4. Estimation Results

In this section I conduct a series of estimations with the goal of measuring the impact of improved inter-provincial connectivity on firm size dispersion in Turkish manufacturing industries. I estimate the equation (4.1) in time differences (midpoint formula), at the province-industry level. I include province and industry fixed effects in all specifications ($\alpha_i$ and $\alpha_s$, respectively).

Since $\Delta DemandConnectivity_i$ varies at the province level, it will be subsumed into the province fixed effect if separately included into the equation, thus rendering the estimation of the coefficient impossible. Therefore, I include $\Delta DemandConnectivity_i$ as a control by interacting it with industry-specific characteristics. Similarly, since industry characteristics (*i.e.* average log weight-to-value ratio, time sensitivity, and share of expenditure on transportation) vary at the same level with the industry fixed effects, I only include them into the estimations by interacting them with all other variables. All observations are weighted by (the inverse of) total employment in province-industry pairs. Robust standard errors are clustered at the province–industry level.

#### 4.4.1. Baseline Results

Table 4.1 lists the baseline estimation results for equation (4.1). Column (1) reports the coefficient estimates where the change in the coefficient of variation measure of firm size dispersion that is calculated based on number of employees of the firm, while column (2) does the same while measuring the firm size dispersion based on firm sales, excluding controls. The estimated coefficient of the two principal variables ($\Delta Competition^{output}$ and $\Delta Competition^{input}$) are negative and statistically significant. This finding indicates that the larger the exposure of industries to the producers of the same industry located elsewhere, the smaller the firm size dispersion in the said industry. Also, industries with increased exposure to potential supplier industries in other provinces (as the result of the infrastructure program), have reduced firm size dispersion. Measuring connectivity terms of weighted averages makes direct quantitative interpretation the coefficients difficult. However, the results constitute a solid evidence toward the directional impact of the infrastructure upgrades on firm size dispersion.

These results might be statistically misleading because of the amount of heterogeneity across industries. To address this issue I include industry characteristics as controls, by interacting them with the connectivity measures. Column (3) of table 4.1 reports the estimation results for equation (4.1), where the dependent variable is the change in the coefficient of variation measure of firm size dispersion that is calculated based on number of employees of the firm, while column (4) does the same while measuring the firm size dispersion based on firm sales. When properly controlled for, the coefficient estimate of $\Delta Competition^{output}$ ceases to be statistically significant, while the coefficient for $\Delta Competition^{input}$ remains significant and within the plausible range (1 standard error) of the "uncontrolled" coefficient in terms of magnitude. The empirical evidence points towards the fact that firm size dispersion, *i.e.* efficiency loss decreases with more access to potential suppliers. A likely mechanism in play here is that, less connected markets induce high trade costs, that only large firms can bear. Hence, decreases in costs of trading across remote markets allow smaller firms have access to cheaper inputs, reducing the firm size dispersion.

Table 4.1: Baseline estimation results of the effect of connectivity on firm size dispersion.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | $\Delta CV$ employment | $\Delta CV$ sales | $\Delta CV$ employment | $\Delta CV$ sales |
| $\Delta Competition^{output}$ | -11.247*** | -12.579*** | 3.018 | 3.948 |
| | (4.014) | (5.73) | (2.749) | (4.054) |
| $\Delta Competition^{input}$ | -19.563*** | -18.221*** | -22.312*** | -20.889** |
| | (6.654) | (6.554) | (7.860) | (7.832) |
| Controls | NO | NO | YES | YES |
| Province FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Observations | 1125 | 1125 | 1125 | 1125 |
| Adj. R-squared | 0.269 | 0.293 | 0.436 | 0.503 |

*Notes:* "$\Delta CV$ employment" refers to the midpoint formula change in coefficient of variation of firm sizes by employment. "$\Delta CV$ sales" refers to the midpoint formula change in coefficient of variation of firm sizes by sales. Industry characteristics, *i.e.* average log weight-to-value ratio, time sensitivity, and share of expenditure on transportation are interacted with connectivity measures in columns (3) and (4). All observations are weighted by total employment in province-industry pairs. Robust standard errors are clustered at the province–industry level. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively.

### 4.4.2. Robustness Checks

In this section I conduct a series of robustness checks to challenge the baseline model to misspecification. Additional tables for robustness checks can be found in the appendix of this chapter (tables B5 - B11).

*A: Dispersion measured in Theil index.* First, I start by using an alternative method to calculate firm size dispersion in markets. I use Theil index (equation (4.3)) an re-estimate the baseline model given. Columns (1) and (2) of table 4.2 report coefficient estimates for equation (4.1) without controls for the change in employment and sales dispersion, respectively. The estimates are negative and statistically significant. Columns (3) and (4) report the results for the estimation with full set of controls. The coefficient for $\Delta Competition^{output}$ is not significant, while the coefficient estimates for $\Delta Competition^{input}$ are. This confirms the findings in the baseline results that, increased exposure to potential suppliers in other provinces reduce firm size dispersion in local industries.

Table 4.2: Robustness Check I: the effect of connectivity on firm size dispersion - dispersion calculated as Theil index.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | $\Delta Thl$ employment | $\Delta Thl$ sales | $\Delta Thl$ employment | $\Delta Thl$ sales |
| $\Delta Competition^{output}$ | -9.991** | -9.998** | 2.782 | 3.354 |
| | (4.946) | (4.806) | (2.739) | (3.117) |
| $\Delta Competition^{input}$ | -16.004** | -18.859*** | -17.747** | -20.190*** |
| | (7.767) | (4.714) | (7.589) | (7.133) |
| Controls | NO | NO | YES | YES |
| Province FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Observations | 1125 | 1125 | 1125 | 1125 |
| Adj. R-squared | 0.331 | 0.324 | 0.390 | 0.459 |

*Notes:* "$\Delta Thl$ employment" refers to the midpoint formula change in Theil index measure of firm sizes by employment. "$\Delta Thl$ sales" refers to the midpoint formula change in Theil index measure of firm sizes by sales. Industry characteristics, *i.e.* average log weight-to-value ratio, time sensitivity, and share of expenditure on transportation are interacted with connectivity measures in columns (3) and (4). All observations are weighted by total employment in province-industry pairs. Robust standard errors are clustered at the province–industry level. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively.

*B: Changes in firm size dispersion measured in logarithmic differences.* Alternative to the baseline model, where the changes are calculated via the midpoint formula (equation (4.4)), I now calculate change in time in logarithmic difference terms. Table 4.3 summarizes the coefficient estimates for equation (4.1).

Once again, the behavior of coefficient estimates are in line with the baseline specification. When estimated without controls, both connectivity measures have negative, statistically significant effect on firm size dispersion (columns (1) and (2)), whereas when properly controlled for industry heterogeneity, only the exposure to potential suppliers in other provinces has a significant effect on reducing firm size dispersion.

Table 4.3: Robustness Check II: the effect of connectivity on firm size dispersion - time change calculated logarithmic difference.

| | (1) $\Delta CV$ employment | (2) $\Delta CV$ sales | (3) $\Delta CV$ employment | (4) $\Delta CV$ sales |
|---|---|---|---|---|
| $\Delta Competition^{output}$ | -10.567*** | -12.639*** | 4.280 | 5.328 |
| | (3.645) | (4.51) | (3.218) | (4.402) |
| $\Delta Competition^{input}$ | -18.436*** | -16.869*** | -22.259** | -20.683** |
| | (6.584) | (5.917) | (8.795) | (8.491) |
| Controls | NO | NO | YES | YES |
| Province FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Observations | 956 | 962 | 937 | 946 |
| Adj. R-squared | 0.290 | 0.265 | 0.446 | 0.440 |

*Notes:* "$\Delta CV$ employment" refers to the logarithmic change in coefficient of variation of firm sizes by employment. "$\Delta CV$ sales" refers to the logarithmic change in coefficient of variation of firm sizes by sales. Industry characteristics, *i.e.* average log weight-to-value ratio, time sensitivity, and share of expenditure on transportation are interacted with connectivity measures in columns (3) and (4). All observations are weighted by total employment in province-industry pairs. Robust standard errors are clustered at the province–industry level. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively.

*C: Regressions with alternative analytical weight.* One of the useful econometric tools is weighting observations. The general purpose of this procedure is to eliminate potential heteroskedastic behavior in the error term. In our case, province-industry pairs vary widely in terms of the volume of production and potential for trade. It is plausible to expect that observations from large province-industries (such as auto industry in Bursa) have different error term variance than the smaller ones (such as chemical industry in Artvin). As a result, standard error estimates are prone to be faulty, rendering statistical inference faulty as well. Weighting each observation by the inverse square root of a common variable that likely explains the differences in error term variance eliminates these problems. I chose the total employment in province-industries in 2006 as weights in the baseline specification. As a robustness check, I re-estimate the baseline model by weighting each observation by total value of sales in 2006. Columns (1) and (2) of table 4.4 report coefficient estimates for equation (4.1) without controls for the change in employment and sales dispersion, respectively. The estimates are negative and statistically significant. Columns (3) and (4) report the results for the estimation with full set of controls. The coefficient for $\Delta Competition^{output}$ is not significant, while the coefficient estimates for $\Delta Competition^{input}$ are, once again confirming the findings in the baseline results that, increased exposure to potential suppliers in other provinces reduce firm size dispersion in local industries.

Table 4.4: Robustness Check III: the effect of connectivity on firm size dispersion - observations weighted by sales.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | $\Delta CV$ employment | $\Delta CV$ sales | $\Delta CV$ employment | $\Delta CV$ sales |
| $\Delta Competition^{output}$ | -13.943* | -13.386** | 3.427 | 4.815 |
| | (7.11) | (5.33) | (4.065) | (4.493) |
| $\Delta Competition^{input}$ | -16.887* | -17.747** | -15.115* | -18.990** |
| | (8.451) | (7.589) | (7.533) | (7.333) |
| Controls | NO | NO | YES | YES |
| Province FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Observations | 1125 | 1125 | 1125 | 1125 |
| Adj. R-squared | 0.209 | 0.213 | 0.306 | 0.363 |

*Notes:* "$\Delta CV$ employment" refers to the midpoint formula change in coefficient of variation of firm sizes by employment. "$\Delta CV$ sales" refers to the midpoint formula change in coefficient of variation of firm sizes by sales. Industry characteristics, *i.e.* average log weight-to-value ratio, time sensitivity, and share of expenditure on transportation are interacted with connectivity measures in columns (3) and (4). All observations are weighted by total sales in province-industry pairs. Robust standard errors are clustered at the province–industry level. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively.

*D: Regressions excluding Istanbul.* Istanbul is the largest province of Turkey accounting for 18% of the total population and 30% of GDP (cite worldbank). Being a large economic powerhouse, Istanbul can be dominating the results in the baseline specification. As a robustness check, I restrict the sample by excluding

Istanbul from the analysis and re-estimate the model. Table 4.5 summarizes the estimation results. The behavior of the coefficient estimates are similar to the baseline specification in magnitudes.

Table 4.5: Robustness Check IV: the effect of connectivity on firm size dispersion - excluding Istanbul.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | $\Delta CV$ employment | $\Delta CV$ sales | $\Delta CV$ employment | $\Delta CV$ sales |
| $\Delta Competition^{output}$ | -10.953** | -10.965** | 3.096 | 1.718 |
| | (4.464) | (4.287) | (7.308) | (6.670) |
| $\Delta Competition^{input}$ | -18.722** | -19.034*** | -17.716* | -23.331** |
| | (7.512) | (6.398) | (8.784) | (8.334) |
| Controls | NO | NO | YES | YES |
| Province FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Observations | 1100 | 1100 | 1100 | 1100 |
| Adj. R-squared | 0.323 | 0.269 | 0.354 | 0.261 |

*Notes:* "$\Delta CV$ employment" refers to the midpoint formula change in coefficient of variation of firm sizes by employment. "$\Delta CV$ sales" refers to the midpoint formula change in coefficient of variation of firm sizes by sales. Industry characteristics, *i.e.* average log weight-to-value ratio, time sensitivity, and share of expenditure on transportation are interacted with connectivity measures in columns (3) and (4). All observations are weighted by total employment in province-industry pairs. Robust standard errors are clustered at the province–industry level. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively.

## 4.5. Conclusion

In this study, I utilize Turkey's extensive infrastructure development program from 2005 to 2015 to examine how enhanced connectivity among Turkish provinces impacts firm size dispersion within local industries. I create variables to separate the effects of direct competition from the effects of access to potential supplier industries. My findings indicate that firm size dispersion, representing efficiency losses, decreases with greater access to potential suppliers. However, there is no evidence that the pro-competitive pressure from increased connectivity influences firm size dispersion.

# CHAPTER 5

# CONCLUSION

This dissertation explores the different ways in which geography shapes economic outcomes, contributing to the field of economic geography and quantitative spatial modeling through three distinct essays.

The first chapter delves into the influence of local markets for non-tradable inputs on urban land prices. By employing a two-stage estimation approach with market access as an instrumental variable, the analysis reveals a positive correlation between the presence of professional businesses and land prices. This relationship, robust to various specifications, highlights the importance of non-tradable inputs in urban economic dynamics and provides a nuanced understanding of the factors driving land price variations.

In the second chapter, the focus shifts to estimating the trade elasticity for Turkish exports at the product level. Through detailed analysis of import tariffs and trade flows, product-specific trade elasticities were determined, revealing important insights into the characteristics of Turkish products and the gains from trade. The findings showed that trade elasticity varies by product type, with homogeneous goods and consumption goods exhibiting slightly different elasticities compared to differentiated and capital goods. The calculated welfare gains from trade, both homogeneous and heterogeneous, underscore the significant benefits of international trade for the Turkish economy.

The third chapter investigates the effects of Turkey's extensive infrastructure development program on the size dispersion within local industries. The results indicate that improved connectivity reduces size dispersion by enhancing access to potential suppliers, while I find no evidence of the pro-competitive pressure from increased connectivity on size dispersion. This effect suggests that reduced trade costs benefit smaller firms, allowing them to compete more effectively.

Overall, this dissertation provides valuable insights into the interplay between geography and economic activity. By addressing key gaps in the literature, it enhances our understanding of urban economics, international trade, and domestic trade. The findings have important implications for policymakers, highlighting the need for strategic infrastructure investments and the consideration of geographical factors in economic planning. As the global economy continues to evolve, further research in economic geography will be essential for uncovering new dimensions of how physical space influences economic opportunities and outcomes.

# BIBLIOGRAPHY


Aghion, P., Bloom, N., Blundell, R., Griffith, R. & Howitt, P. (2005), 'Competition and innovation: An inverted-u relationship', *The Quarterly Journal of Economics* **120**(2), 701–728.

Allen, T. & Arkolakis, C. (2014), 'Trade and the Topography of the Spatial Economy', *The Quarterly Journal of Economics* **129**(3), 1085–1140.
**URL:** *https://ideas.repec.org/a/oup/qjecon/v129y2014i3p1085-1140.html*

Alonso, W. (1964), 'The historic and the structural theories of urban form: Their implications for urban renewal', *Land economics* **40**(2), 227–231.

Anderson, J. E. (1979), 'A theoretical foundation for the gravity equation', *The American economic review* **69**(1), 106–116.

Anderson, J. E. & Van Wincoop, E. (2003), 'Gravity with gravitas: A solution to the border puzzle', *American economic review* **93**(1), 170–192.

Arkolakis, C., Costinot, A. & Rodríguez-Clare, A. (2012), 'New trade models, same old gains?', *American Economic Review* **102**(1), 94–130.
**URL:** *https://www.aeaweb.org/articles?id=10.1257/aer.102.1.94*

Asturias, J., García-Santana, M. & Ramos, R. (2018), 'Competition and the Welfare Gains from Transportation Infrastructure: Evidence from the Golden Quadrilateral of India', *Journal of the European Economic Association*

**17**(6), 1881–1940.
**URL:** *https://doi.org/10.1093/jeea/jvy039*

Bas, M. & Strauss-Kahn, V. (2015), 'Input-trade liberalization, export prices and quality upgrading', *Journal of International Economics* **95**(2), 250–262.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0022199614001500*

Behrens, K., Mion, G., Murata, Y. & Suedekum, J. (2017), 'Spatial frictions', *Journal of Urban Economics* **97**(C), 40–70.
**URL:** *https://ideas.repec.org/a/eee/juecon/v97y2017icp40-70.html*

Black, D., McKinnish, T. & Sanders, S. (2005), 'The Economic Impact Of The Coal Boom And Bust', *Economic Journal* **115**(503), 449–476.
**URL:** *https://ideas.repec.org/a/ecj/econjl/v115y2005i503p449-476.html*

Bolatto, S. & Moramarco, G. (2023), 'Gains from trade and their quantification: Does sectoral disaggregation matter?', *International Economics* **174**, 44–68.
**URL:** *https://www.sciencedirect.com/science/article/pii/S2110701723000203*

Caliendo, L. & Parro, F. (2015), 'Estimates of the trade and welfare effects of nafta', *The Review of Economic Studies* **82**(1), 1–44.

Combes, P.-P., Duranton, G. & Gobillon, L. (2018), 'The Costs of Agglomeration: House and Land Prices in French Cities', *The Review of Economic Studies* **86**(4), 1556–1589.
**URL:** *https://doi.org/10.1093/restud/rdy063*

Combes, P.-P., Duranton, G., Gobillon, L., Puga, D. & Roux, S. (2012), 'The productivity advantages of large cities: Distinguishing agglomeration from firm selection', *Econometrica* **80**(6), 2543–2594.

Costinot, A. & Rodríguez-Clare, A. (2014), *Chapter 4 - Trade Theory with Numbers: Quantifying the Consequences of Globalization*, Vol. 4 of *Handbook of*

*International Economics*, Elsevier.
**URL:** *https://www.sciencedirect.com/science/article/pii/B9780444543141000045*

Coşar, A. K. & Demir, B. (2016), 'Domestic road infrastructure and international trade: Evidence from turkey', *Journal of Development Economics* **118**, 232–244.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0304387815001121*

Coşar, A. K., Demir, B., Ghose, D. & Young, N. (2021*a*), 'Road capacity, domestic trade and regional outcomes', *Journal of Economic Geography* **22**(5), 901–929.
**URL:** *https://doi.org/10.1093/jeg/lbab045*

Coşar, A. K., Demir, B., Ghose, D. & Young, N. (2021*b*), 'Road capacity, domestic trade and regional outcomes', *Journal of Economic Geography* **22**(5), 901–929.

Davis, S. J. & Haltiwanger, J. (1998), Measuring Gross Worker and Job Flows, *in* 'Labor Statistics Measurement Issues', NBER Chapters, National Bureau of Economic Research, Inc, pp. 77–122.
**URL:** *https://ideas.repec.org/h/nbr/nberch/8358.html*

De Loecker, J., Eeckhout, J. & Unger, G. (2020), 'The rise of market power and the macroeconomic implications', *The Quarterly Journal of Economics* **135**(2), 561–644.

Demir, B., Fieler, A. C., Xu, D. Y. & Yang, K. K. (2024), 'O-ring production networks', *Journal of Political Economy* **132**(1), 200–247.

Donaldson, D. (2018), 'Railroads of the raj: Estimating the impact of transportation infrastructure', *American Economic Review* **108**(4-5), 899–934.

Donaldson, D. & Hornbeck, R. (2016), ' Railroads and American Economic Growth: A "Market Access" Approach *', *The Quarterly Journal of Economics*

**131**(2), 799–858.
**URL:** *https://doi.org/10.1093/qje/qjw002*

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

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

Edmond, C., Midrigan, V. & Xu, D. Y. (2015), 'Competition, markups, and the gains from international trade', *American Economic Review* **105**(10), 3183–3221.
**URL:** *https://www.aeaweb.org/articles?id=10.1257/aer.20120549*

Feenstra, R. C. (1994), 'New product varieties and the measurement of international prices', *The American Economic Review* pp. 157–177.

Feng, L., Li, Z. & Swenson, D. L. (2016), 'The connection between imported intermediate inputs and exports: Evidence from chinese firms', *Journal of International Economics* **101**, 86–101.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0022199616300393*

Fontagné, L., Guimbard, H. & Orefice, G. (2022*a*), 'A new dataset on product-level trade elasticities', *Data in Brief* **45**, 108668.
**URL:** *https://www.sciencedirect.com/science/article/pii/S2352340922008733*

Fontagné, L., Guimbard, H. & Orefice, G. (2022*b*), 'Tariff-based product-level trade elasticities', *Journal of International Economics* **137**, 103593.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0022199622000253*

Fujita, M. & Ogawa, H. (1982), 'Multiple equilibria and structural transition of non-monocentric urban configurations', *Regional Science and Urban Economics* **12**(2), 161–196.

Fujita, M. & Thisse, J.-F. (1996), 'Economics of agglomeration', *Journal of the Japanese and international economies* **10**(4), 339–378.

Giri, R., Yi, K.-M. & Yilmazkuday, H. (2021), 'Gains from trade: Does sectoral heterogeneity matter?', *Journal of International Economics* **129**, 103429.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A. & Shleifer, A. (1992), 'Growth in cities', *Journal of Political Rconomy* **100**(6), 1126–1152.

Harris, C. D. (1954), 'The market as a factor in the localization of industry in the united states', *Annals of the Association of American Geographers* **44**(4), 315–348.
**URL:** *http://www.jstor.org/stable/2561395*

Hornbeck, R. & Rotemberg, M. (2019), Railroads, reallocation, and the rise of american manufacturing, Technical report, National Bureau of Economic Research.

Hsieh, C.-T. & Klenow, P. J. (2009), 'Misallocation and Manufacturing TFP in China and India*', *The Quarterly Journal of Economics* **124**(4), 1403–1448.
**URL:** *https://doi.org/10.1162/qjec.2009.124.4.1403*

Krugman, P. (1991), 'Increasing returns and economic geography', *Journal of political economy* **99**(3), 483–499.

Krugman, P. (1992), *Geography and trade*, MIT press.

Krugman, P. (1996), *The self organizing economy*, John Wiley & Sons.

Krugman, P. R. (1997), *Development, geography, and economic theory*, Vol. 6, MIT press.

Krugman, P. et al. (1980), 'Scale economies, product differentiation, and the pattern of trade', *American economic review* **70**(5), 950–959.

Liu, Q. & Qiu, L. D. (2016), 'Intermediate input imports and innovations: Evidence from chinese firms' patent filings', *Journal of International Economics* **103**, 166–183.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0022199616301131*

Marshall, A. (1890), *The Principles of Economics*, London: Macmillan.

Melitz, M. J. (2003), 'The impact of trade on intra-industry reallocations and aggregate industry productivity', *Econometrica* **71**(6), 1695–1725.

Mills, E. S. (1967), 'An aggregative model of resource allocation in a metropolitan area', *The American Economic Review* **57**(2), 197–210.

Moretti, E. (2010), 'Local Multipliers', *American Economic Review* **100**(2), 373–377.
**URL:** *https://ideas.repec.org/a/aea/aecrev/v100y2010i2p373-77.html*

Moretti, E. (2011), 'Chapter 14 - local labor markets', *Handbook of Labor Economics* **4**, 1237–1313.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169721811024129*

Muth, R. F. (1971), 'The derived demand for urban residential land', *Urban Studies* **8**(3), 243–254.

Ossa, R. (2015), 'Why trade matters after all', *Journal of International Economics* **97**(2), 266–277.

Redding, S. J. & Sturm, D. M. (2008), 'The costs of remoteness: Evidence from german division and reunification', *American Economic Review* **98**(5), 1766–1797.

Redding, S. & Venables, A. J. (2004), 'Economic geography and international inequality', *Journal of International Economics* **62**(1), 53–82.

Silva, J. M. C. S. & Tenreyro, S. (2006), 'The Log of Gravity', *The Review of Economics and Statistics* **88**(4), 641–658.
**URL:** *https://doi.org/10.1162/rest.88.4.641*

Simonovska, I. & Waugh, M. E. (2014*a*), 'The elasticity of trade: Estimates and evidence', *Journal of International Economics* **92**(1), 34–50.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0022199613000986*

Simonovska, I. & Waugh, M. E. (2014*b*), 'The elasticity of trade: Estimates and evidence', *Journal of International Economics* **92**(1), 34–50.

Wu, M., Yu, L. & Zhang, J. (2023), 'Road expansion, allocative efficiency, and pro-competitive effect of transport infrastructure: Evidence from china', *Journal of Development Economics* **162**, 103050.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0304387823000056*

# APPENDIX

Following tables supplement the results discussed in relevant sections. Tables A3 and A4 are robustness checks for the analyses in chapter 1. Tables B5-B11 report additional robustness checks for analyses in chapter 3.

Table A1: Impact of Office Propensity (measured as the fraction of housing units with business licenses within all housing units) on Land Prices (Varying Trade Elasticities)

| Dependent Variable: | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Log Land Prices | OLS | 2SLS | 2SLS | 2SLS |
| Log Office Propensity | 0.712** | 1.634*** | 1.108* | 1.763*** |
| | (0.321) | (0.383) | (0.588) | (0.298) |
| First-Stage Coefficient | | 0.336** | 1.624** | 0.185** |
| | | (0.129) | (0.740) | (0.072) |
| First-Stage F-statistic | | 24.53 | 24.41 | 21.48 |
| Observations | 953 | 953 | 953 | 953 |
| Province Fixed Effects | YES | YES | YES | YES |
| Geographical Controls | YES | YES | YES | YES |
| Socio-Economic Controls | YES | YES | YES | YES |

*Notes:* Robust standard errors are clustered at the province level and are reported in parentheses. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively. The detailed list of control variables can be found in the first paragraph of Section 4. Columns (2), (3), and (4) include $\beta$ estimates, where market access in equation 2.2 is calculated with $\theta = 4.14$; $\theta = 1$ and $\theta = 8.22$, respectively, excluding all districts within a 100-km radius.

Table A2: Impact of Office Propensity (measured as the fraction of housing units with business licenses within all housing units) on Land Prices (Varying Exclusion Zones)

| Dependent Variable: | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Log Land Prices | OLS | 2SLS | 2SLS | 2SLS |
| Log Office Propensity | 0.712** | 1.634*** | 1.381*** | 1.514* |
| | (0.321) | (0.383) | (0.413) | (0.816) |
| First-Stage Coefficient | | 0.336** | 0.179** | 0.224*** |
| | | (0.129) | (0.083) | (0.079) |
| First-Stage F-statistic | | 24.53 | 26.20 | 21.20 |
| Observations | 953 | 953 | 953 | 953 |
| Province Fixed Effects | YES | YES | YES | YES |
| Geographical Controls | YES | YES | YES | YES |
| Socio-Economic Controls | YES | YES | YES | YES |

*Notes:* Robust standard errors are clustered at the province level and are reported in parentheses. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively. The detailed list of control variables can be found in the first paragraph of Section 4. Columns (2), (3), and (4) include $\beta$ estimates, where market access in equation 2.2 is calculated with $\theta = 4.14$, excluding all districts within a 100-km, 75-km, and 125-km radius, respectively.

Table A3: Impact of Office Propensity (measured as the fraction of housing units with business licenses within all housing units) on Land Prices (Varying Trade Elasticities), including control variables.

| Dependent Variable: Log Land Prices | (1) OLS | (2) 2SLS | (3) 2SLS | (4) 2SLS |
|---|---|---|---|---|
| Log Office Propensity | 0.587** | 1.243*** | 0.855** | 1.323*** |
| | (0.240) | (0.225) | (0.390) | (0.200) |
| IHS Distance to Center | -0.133*** | -0.106*** | -0.122*** | -0.103*** |
| | (0.020) | (0.034) | (0.031) | (0.033) |
| Log District Area ($km^2$) | -0.152*** | -0.053 | -0.112* | -0.046 |
| | (0.046) | (0.054) | (0.062) | (0.052) |
| IHS Distance to Port | -0.020 | 0.026 | -0.001 | 0.030 |
| | (0.026) | (0.033) | (0.032) | (0.035) |
| IHS Distance to Border | 0.140 | 0.118 | 0.131 | 0.117 |
| | (0.103) | (0.099) | (0.100) | (0.101) |
| Coast (Dummy) | 0.078 | 0.035 | 0.061 | 0.031 |
| | (0.069) | (0.073) | (0.075) | (0.073) |
| SEDI | 0.154*** | 0.158*** | 0.156*** | 0.158*** |
| | (0.037) | (0.034) | (0.035) | (0.034) |
| WAP Share (2014) | 5.569* | 5.486*** | 5.535** | 5.480*** |
| | (2.971) | (1.571) | (2.222) | (1.555) |
| College Grad. Share (2014) | 2.055 | 2.159* | 2.097* | 2.167* |
| | (1.363) | (1.154) | (1.250) | (1.146) |
| First-Stage Coefficient | | 0.441** | 2.105* | 0.247* |
| | | (0.217) | (1.167) | (0.126) |
| First-Stage F-statistic | | 32.31 | 35.05 | 27.74 |
| Observations | 953 | 953 | 953 | 953 |
| Province Fixed Effects | YES | YES | YES | YES |
| Geographical Controls | YES | YES | YES | YES |
| Socio-Economic Controls | YES | YES | YES | YES |

*Notes:* Robust standard errors are clustered at the province level and are reported in parentheses. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively. IHS refers to hyperbolic sine transformation. Distances are closest distances to units. Columns (2), (3), and (4) include $\beta$ estimates, where market access in equation 2.2 is calculated with $\theta = 4.14$; $\theta = 1$ and $\theta = 8.22$, respectively, excluding all districts within a 100-km radius.

Table A4: Impact of Office Propensity (measured as the fraction of housing units with business licenses within all housing units) on Land Prices (Varying Exclusion Zones), including control variables.

| Dependent Variable: Log Land Prices | (1) OLS | (2) 2SLS | (3) 2SLS | (4) 2SLS |
|---|---|---|---|---|
| Log Office Propensity | 0.587** | 1.243*** | 1.034*** | 1.108** |
| | (0.240) | (0.225) | (0.310) | (0.540) |
| IHS Distance to Center | -0.133*** | -0.106*** | -0.114*** | -0.111*** |
| | (0.020) | (0.034) | (0.028) | (0.037) |
| Log District Area ($km^2$) | -0.152*** | -0.053 | -0.085 | -0.074 |
| | (0.046) | (0.054) | (0.065) | (0.079) |
| IHS Distance to Port | -0.020 | 0.026 | 0.011 | 0.016 |
| | (0.026) | (0.033) | (0.028) | (0.042) |
| IHS Distance to Border | 0.140 | 0.118 | 0.125 | 0.123 |
| | (0.103) | (0.099) | (0.101) | (0.097) |
| Coast (Dummy) | 0.078 | 0.035 | 0.049 | 0.044 |
| | (0.069) | (0.073) | (0.079) | (0.078) |
| SEDI | 0.154*** | 0.158*** | 0.157*** | 0.157*** |
| | (0.037) | (0.034) | (0.035) | (0.034) |
| WAP Share (2014) | 5.569* | 5.486*** | 5.513*** | 5.503*** |
| | (2.971) | (1.571) | (1.854) | (1.695) |
| College Grad. Share (2014) | 2.055 | 2.159* | 2.126* | 2.138* |
| | (1.363) | (1.154) | (1.207) | (1.169) |
| First-Stage Coefficient | | 0.441** | 0.239* | 0.306*** |
| | | (0.217) | (0.132) | (0.099) |
| First-Stage F-statistic | | 32.31 | 48.63 | 37.21 |
| Observations | 953 | 953 | 953 | 953 |
| Province Fixed Effects | YES | YES | YES | YES |
| Geographical Controls | YES | YES | YES | YES |
| Socio-Economic Controls | YES | YES | YES | YES |

*Notes:* Robust standard errors are clustered at the province level and are reported in parentheses. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively. IHS refers to hyperbolic sine transformation. Distances are closest distances to units. Columns (2), (3), and (4) include $\beta$ estimates, where market access in equation 2.2 is calculated with $\theta = 4.14$, excluding all districts within a 100-km, 75-km, and 125-km radius, respectively.

Table B1: Robustness Check V: the effect of connectivity on size dispersion - dispersion calculated as Theil index, time change calculated logarithmic difference.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | $\Delta Thl$ emp | $\Delta Thl$ sales | $\Delta Thl$ emp | $\Delta Thl$ sales |
| $\Delta Competition^{output}$ | -9.531** | -10.10*** | 4.014 | 4.815 |
| | (4.683) | (3.367) | (2.977) | (3.317) |
| $\Delta Competition^{input}$ | -15.567** | -16.259*** | -16.887* | -20.190*** |
| | (7.646) | (4.279) | (8.451) | (7.133) |
| Controls | NO | NO | YES | YES |
| Province FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Observations | 956 | 962 | 937 | 946 |
| Adj. R-squared | 0.334 | 0.326 | 0.440 | 0.432 |

*Notes:* "$\Delta Thl$ emp" refers to the the logarithmic change in Theil index measure of firm sizes by employment. "$\Delta Thl$ sales" refers to the the logarithmic change in Theil index measure of firm sizes by sales. Industry characteristics, *i.e.* average log weight-to-value ratio, time sensitivity, and share of expenditure on transportation are interacted with connectivity measures in columns (3) and (4). All observations are weighted by total employment in province-industry pairs. Robust standard errors are clustered at the province–industry level. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively.

Table B2: Robustness Check VI: the effect of connectivity on size dispersion - dispersion calculated as Theil index, observations weighted by sales.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | $\Delta Thl$ emp | $\Delta Thl$ sales | $\Delta Thl$ emp | $\Delta Thl$ sales |
| $\Delta Competition^{output}$ | -14.690** | -14.120** | 2.157 | 2.812 |
| | (5.852) | (5.325) | (3.727) | (2.412) |
| $\Delta Competition^{input}$ | -20.683** | -20.889** | -7.366 | -11.862* |
| | (8.491) | (7.832) | (7.924) | (5.976) |
| Controls | NO | NO | YES | YES |
| Province FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Observations | 1125 | 1125 | 1125 | 1125 |
| Adj. R-squared | 0.246 | 0.254 | 0.312 | 0.371 |

*Notes:* "$\Delta Thl$ emp" refers to the midpoint formula change in Theil index measure of firm sizes by employment. "$\Delta Thl$ sales" refers to the midpoint formula change in Theil index measure of firm sizes by sales. Industry characteristics, *i.e.* average log weight-to-value ratio, time sensitivity, and share of expenditure on transportation are interacted with connectivity measures in columns (3) and (4). All observations are weighted by total sales in province-industry pairs. Robust standard errors are clustered at the province–industry level. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively.

Table B3: Robustness Check VII: the effect of connectivity on size dispersion - dispersion calculated as Theil index, excluding Istanbul.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | $\Delta Thl$ emp | $\Delta Thl$ sales | $\Delta Thl$ emp | $\Delta Thl$ sales |
| $\Delta Competition^{output}$ | -11.703* | -11.629* | 4.294 | 4.976 |
| | (5.726) | (6.013) | (5.709) | (5.466) |
| $\Delta Competition^{input}$ | -21.906** | -21.930** | -16.963** | -17.981** |
| | (8.928) | (8.574) | (7.615) | (6.411) |
| Controls | NO | NO | YES | YES |
| Province FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Observations | 1100 | 1100 | 1100 | 1100 |
| Adj. R-squared | 0.313 | 0.296 | 0.340 | 0.345 |

*Notes:* "$\Delta Thl$ emp" refers to the midpoint formula change in Theil index measure of firm sizes by employment. "$\Delta Thl$ sales" refers to the midpoint formula change in Theil index measure of firm sizes by sales. Industry characteristics, *i.e.* average log weight-to-value ratio, time sensitivity, and share of expenditure on transportation are interacted with connectivity measures in columns (3) and (4). All observations are weighted by total employment in province-industry pairs. Robust standard errors are clustered at the province–industry level. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively.

Table B4: The effect of connectivity on size dispersion - including control variables

| Dependent variable: d size dispersion | (1) $\Delta CV$ emp | (2) $\Delta CV$ sales | (3) $\Delta CV$ emp | (4) $\Delta CV$ sales |
|---|---|---|---|---|
| $\Delta Competition^{output}$ | -11.247*** | -12.579*** | 3.018 | 3.948 |
| | (4.014) | (5.73) | (2.749) | (4.054) |
| $\Delta Competition^{input}$ | -19.563*** | -18.221*** | -22.312*** | -20.889** |
| | (6.654) | (6.554) | (7.860) | (7.832) |
| $\Delta Competition^{output} * Heavy$ | | | 0.790 | 2.903 |
| | | | (1.000) | (2.704) |
| $\Delta Competition^{input} * Heavy$ | | | -6.524** | -7.859 |
| | | | (2.643) | (4.886) |
| $\Delta DemandConn * Heavy$ | | | 4.995 | 4.044 |
| | | | (3.728) | (4.655) |
| $\Delta Competition^{output} * TS$ | | | -7.858 | -7.960 |
| | | | (5.742) | (5.511) |
| $\Delta Competition^{input} * TS$ | | | 17.806 | 11.711 |
| | | | (11.157) | (14.594) |
| $\Delta DemandConn * TS$ | | | -23.174 | -18.252 |
| | | | (17.330) | (16.994) |
| $\Delta Competition^{output} * Sh.Transp.$ | | | -520.568** | -742.690* |
| | | | (189.354) | (365.074) |
| $\Delta Competition^{input} * Sh.Transp.$ | | | 1,629.309*** | 1,484.225** |
| | | | (379.029) | (628.223) |
| $\Delta DemandConn * Sh.Transp.$ | | | -1,118.391** | -700.330 |
| | | | (531.196) | (583.585) |
| Controls | NO | NO | YES | YES |
| Province FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Observations | 1,125 | 1,125 | 1,125 | 1,125 |
| Adj. R-squared | 0.269 | 0.293 | 0.436 | 0.503 |

*Notes:* "$\Delta CV$ emp" refers to the the logarithmic change in CV index measure of firm sizes by employment. "$\Delta CV$ sales" refers to the the logarithmic change in CV index measure of firm sizes by sales. Industry characteristics, *i.e.* average log weight-to-value ratio ($Heavy$), time sensitivity ($TS$), and share of expenditure on transportation ($Sh.Transp.$) are interacted with connectivity measures in columns (3) and (4). All observations are weighted by total employment in province-industry pairs. Robust standard errors are clustered at the province–industry level. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively.

Table B5: The effect of connectivity on size dispersion - dispersion calculated as Theil index, including control variables

| Dependent variable: d size dispersion | (1) $\Delta Thl$ emp | (2) $\Delta Thl$ sales | (3) $\Delta Thl$ emp | (4) $\Delta Thl$ sales |
|---|---|---|---|---|
| $\Delta Competition^{output}$ | -9.991** | -9.998** | 2.782 | 3.354 |
| | (4.946) | (4.806) | (2.739) | (3.117) |
| $\Delta Competition^{input}$ | -16.004** | -18.859*** | -17.747** | -20.190*** |
| | (7.767) | (4.714) | (7.589) | (7.133) |
| $\Delta Competition^{output}$ * $Heavy$ | | | -0.100 | 0.613 |
| | | | (1.462) | (1.435) |
| $\Delta Competition^{input}$ * $Heavy$ | | | -3.973 | -4.112 |
| | | | (3.292) | (3.422) |
| $\Delta DemandConn$ * $Heavy$ | | | 2.689 | 2.499 |
| | | | (4.554) | (3.751) |
| $\Delta Competition^{output}$ * $TS$ | | | -10.160 | -7.045 |
| | | | (6.040) | (5.060) |
| $\Delta Competition^{input}$ * $TS$ | | | 17.904 | 16.515 |
| | | | (10.932) | (9.758) |
| $\Delta DemandConn$ * $TS$ | | | -19.289 | -21.221 |
| | | | (16.815) | (13.881) |
| $\Delta Competition^{output}$ * $Sh.Transp.$ | | | -438.462*** | -479.153** |
| | | | (143.006) | (224.342) |
| $\Delta Competition^{input}$ * $Sh.Transp.$ | | | 1,259.775*** | 1,275.032** |
| | | | (431.806) | (491.181) |
| $\Delta DemandConn$ * $Sh.Transp.$ | | | -783.253 | -785.767* |
| | | | (562.731) | (450.394) |
| Controls | NO | NO | YES | YES |
| Province FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Observations | 1,125 | 1,125 | 1,125 | 1,125 |
| Adj. R-squared | 0.331 | 0.324 | 0.390 | 0.459 |

*Notes:* "$\Delta Thl$ emp" refers to the the logarithmic change in Theil index measure of firm sizes by employment. "$\Delta Thl$ sales" refers to the the logarithmic change in Theil index measure of firm sizes by sales. Industry characteristics, *i.e.* average log weight-to-value ratio ($Heavy$), time sensitivity ($TS$), and share of expenditure on transportation ($Sh.Transp.$) are interacted with connectivity measures in columns (3) and (4). All observations are weighted by total employment in province-industry pairs. Robust standard errors are clustered at the province–industry level. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively.

Table B6: The effect of connectivity on size dispersion - time change calculated as logarithmic differences, including control variables.

| Dependent variable: d size dispersion | (1) $\Delta CV$ emp | (2) $\Delta CV$ sales | (3) $\Delta CV$ emp | (4) $\Delta CV$ sales |
|---|---|---|---|---|
| $\Delta Competition^{output}$ | -10.567*** | -12.639*** | 4.280 | 5.328 |
| | (3.645) | (4.51) | (3.218) | (4.402) |
| $\Delta Competition^{input}$ | -18.436*** | -16.869*** | -22.259** | -20.683** |
| | (6.584) | (5.917) | (8.795) | (8.491) |
| $\Delta Competition^{output}$ * $Heavy$ | | | 1.554 | 3.865 |
| | | | (1.295) | (2.730) |
| $\Delta Competition^{input}$ * $Heavy$ | | | -6.159** | -7.500 |
| | | | (2.829) | (5.038) |
| $\Delta DemandConn$ * $Heavy$ | | | 3.360 | 2.131 |
| | | | (3.806) | (4.709) |
| $\Delta Competition^{output}$ * $TS$ | | | -6.084 | -6.105 |
| | | | (5.582) | (5.203) |
| $\Delta Competition^{input}$ * $TS$ | | | 18.137 | 10.926 |
| | | | (11.951) | (15.025) |
| $\Delta DemandConn$ * $TS$ | | | -26.389 | -20.594 |
| | | | (19.438) | (18.293) |
| $\Delta Competition^{output}$ * $Sh.Transp.$ | | | -641.507*** | -885.580** |
| | | | (202.176) | (398.849) |
| $\Delta Competition^{input}$ * $Sh.Transp.$ | | | 1,607.100*** | 1,456.380** |
| | | | (440.850) | (662.616) |
| $\Delta DemandConn$ * $Sh.Transp.$ | | | -931.117 | -471.353 |
| | | | (567.887) | (598.806) |
| Controls | NO | NO | YES | YES |
| Province FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Observations | 956 | 962 | 937 | 946 |
| Adj. R-squared | 0.29 | 0.265 | 0.446 | 0.440 |

*Notes:* "$\Delta CV$ emp" refers to the midpoint formula change in CV index measure of firm sizes by employment. "$\Delta CV$ sales" refers to the midpoint formula change in CV index measure of firm sizes by sales. Industry characteristics, *i.e.* average log weight-to-value ratio ($Heavy$), time sensitivity ($TS$), and share of expenditure on transportation ($Sh.Transp.$) are interacted with connectivity measures in columns (3) and (4). All observations are weighted by total employment in province-industry pairs. Robust standard errors are clustered at the province–industry level. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively.

Table B7: The effect of connectivity on size dispersion - dispersion calculated as Theil index, time change calculated as logarithmic differences, including control variables.

| Dependent variable: d size dispersion | (1) $\Delta Thl$ emp | (2) $\Delta Thl$ sales | (3) $\Delta Thl$ emp | (4) $\Delta Thl$ sales |
|---|---|---|---|---|
| $\Delta Competition^{output}$ | -9.531** | -10.10*** | 4.014 | 4.815 |
| | (4.683) | (3.367) | (2.977) | (3.317) |
| $\Delta Competition^{input}$ | -15.567** | -16.259*** | -16.887* | -20.190*** |
| | (7.646) | (4.279) | (8.451) | (7.133) |
| $\Delta Competition^{output} * Heavy$ | | | 0.858 | 1.347 |
| | | | (3.251) | (1.351) |
| $\Delta Competition^{input} * Heavy$ | | | -3.123 | -3.757 |
| | | | (3.763) | (3.542) |
| $\Delta DemandConn * Heavy$ | | | 0.364 | 0.812 |
| | | | (5.027) | (3.854) |
| $\Delta Competition^{output} * TS$ | | | -8.727 | -4.569 |
| | | | (6.449) | (4.681) |
| $\Delta Competition^{input} * TS$ | | | 17.846 | 17.020 |
| | | | (11.676) | (10.431) |
| $\Delta DemandConn * TS$ | | | -21.694 | -26.362 |
| | | | (18.824) | (16.890) |
| $\Delta Competition^{output} * Sh.Transp.$ | | | -577.467*** | -606.854** |
| | | | (109.794) | (250.080) |
| $\Delta Competition^{input} * Sh.Transp.$ | | | 1,173.641** | 1,261.145** |
| | | | (510.788) | (536.144) |
| $\Delta DemandConn * Sh.Transp.$ | | | -506.846 | -586.300 |
| | | | (636.284) | (479.418) |
| Controls | NO | NO | YES | YES |
| Province FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Observations | 956 | 962 | 937 | 946 |
| Adj. R-squared | 0.334 | 0.326 | 0.440 | 0.432 |

*Notes:* "$\Delta Thl$ emp" refers to the midpoint formula change in Theil index measure of firm sizes by employment. "$\Delta Thl$ sales" refers to the midpoint formula change in Theil index measure of firm sizes by sales. Industry characteristics, *i.e.* average log weight-to-value ratio ($Heavy$), time sensitivity ($TS$), and share of expenditure on transportation ($Sh.Transp.$) are interacted with connectivity measures in columns (3) and (4). All observations are weighted by total employment in province-industry pairs. Robust standard errors are clustered at the province–industry level. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively.

Table B8: The effect of connectivity on size dispersion - observations weighted by sales, including control variables.

| Dependent variable: d size dispersion | (1) $\Delta CV$ emp | (2) $\Delta CV$ sales | (3) $\Delta CV$ emp | (4) $\Delta CV$ sales |
|---|---|---|---|---|
| $\Delta Competition^{output}$ | -13.943* | -13.386** | 3.427 | 4.815 |
| | (7.11) | (5.33) | (4.065) | (4.493) |
| $\Delta Competition^{input}$ | -16.887* | -17.747** | -15.115* | -18.990** |
| | (8.451) | (7.589) | (7.533) | (7.333) |
| $\Delta Competition^{output}$ * $Heavy$ | | | 1.747 | 3.959 |
| | | | (1.583) | (3.122) |
| $\Delta Competition^{input}$ * $Heavy$ | | | -7.490*** | -9.134* |
| | | | (2.523) | (4.846) |
| $\Delta DemandConn$ * $Heavy$ | | | 5.108 | 3.706 |
| | | | (3.808) | (5.357) |
| $\Delta Competition^{output}$ * $TS$ | | | -3.770 | -3.755 |
| | | | (4.298) | (3.859) |
| $\Delta Competition^{input}$ * $TS$ | | | 6.403 | 1.212 |
| | | | (11.296) | (11.581) |
| $\Delta DemandConn$ * $TS$ | | | -11.344 | -8.349 |
| | | | (16.517) | (14.950) |
| $\Delta Competition^{output}$ * $Sh.Transp.$ | | | -404.919 | -628.214 |
| | | | (249.616) | (419.212) |
| $\Delta Competition^{input}$ * $Sh.Transp.$ | | | 1,339.415*** | 1,342.059* |
| | | | (390.286) | (663.114) |
| $\Delta DemandConn$ * $Sh.Transp.$ | | | -944.699 | -620.071 |
| | | | (569.725) | (727.681) |
| Controls | NO | NO | YES | YES |
| Province FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Observations | 1125 | 1,125 | 1125 | 1,125 |
| Adj. R-squared | 0.209 | 0.213 | 0.306 | 0.363 |

*Notes:* "$\Delta CV$ emp" refers to the midpoint formula change in CV index measure of firm sizes by employment. "$\Delta CV$ sales" refers to the midpoint formula change in CV index measure of firm sizes by sales. Industry characteristics, *i.e.* average log weight-to-value ratio ($Heavy$), time sensitivity ($TS$), and share of expenditure on transportation ($Sh.Transp.$) are interacted with connectivity measures in columns (3) and (4). All observations are weighted by total sales in province-industry pairs. Robust standard errors are clustered at the province–industry level. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively.

Table B9: The effect of connectivity on size dispersion - dispersion calculated as Theil index, observations weighted by sales, including control variables.

| Dependent variable: d size dispersion | (1) $\Delta Thl$ emp | (2) $\Delta Thl$ sales | (3) $\Delta Thl$ emp | (4) $\Delta Thl$ sales |
|---|---|---|---|---|
| $\Delta Competition^{output}$ | -14.690** | -14.120** | 2.157 | 2.812 |
| | (5.852) | (5.325) | (3.727) | (2.412) |
| $\Delta Competition^{input}$ | -20.683** | -20.889** | -7.366 | -11.862* |
| | (8.491) | (7.832) | (7.924) | (5.976) |
| $\Delta Competition^{output} * Heavy$ | | | 0.995 | 2.093 |
| | | | (1.337) | (1.760) |
| $\Delta Competition^{input} * Heavy$ | | | -4.591 | -2.978 |
| | | | (3.284) | (3.639) |
| $\Delta DemandConn * Heavy$ | | | 2.015 | -1.234 |
| | | | (4.712) | (4.329) |
| $\Delta Competition^{output} * TS$ | | | -3.022 | -1.320 |
| | | | (4.416) | (2.663) |
| $\Delta Competition^{input} * TS$ | | | 2.195 | -0.449 |
| | | | (10.663) | (9.459) |
| $\Delta DemandConn * TS$ | | | -6.065 | -6.855 |
| | | | (15.789) | (14.679) |
| $\Delta Competition^{output} * Sh.Transp.$ | | | -251.340 | -402.140* |
| | | | (184.933) | (224.663) |
| $\Delta Competition^{input} * Sh.Transp.$ | | | 837.563* | 732.294 |
| | | | (467.540) | (535.147) |
| $\Delta DemandConn * Sh.Transp.$ | | | -475.596 | -180.925 |
| | | | (615.064) | (589.841) |
| Controls | NO | NO | YES | YES |
| Province FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Observations | 1,125 | 1,125 | 1,125 | 1,125 |
| Adj. R-squared | 0.246 | 0.254 | 0.312 | 0.371 |

*Notes:* "$\Delta Thl$ emp" refers to the midpoint formula change in Theil index measure of firm sizes by employment. "$\Delta Thl$ sales" refers to the midpoint formula change in Theil index measure of firm sizes by sales. Industry characteristics, *i.e.* average log weight-to-value ratio ($Heavy$), time sensitivity ($TS$), and share of expenditure on transportation ($Sh.Transp.$) are interacted with connectivity measures in columns (3) and (4). All observations are weighted by total sales in province-industry pairs. Robust standard errors are clustered at the province–industry level. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively.

Table B10: The effect of connectivity on size dispersion - excluding Istanbul, including control variables.

| Dependent variable: d size dispersion | (1) $\Delta CV$ emp | (2) $\Delta CV$ sales | (3) $\Delta CV$ emp | (4) $\Delta CV$ sales |
|---|---|---|---|---|
| $\Delta Competition^{output}$ | -10.953** | -10.965** | 3.096 | 1.718 |
| | (4.464) | (4.287) | (7.308) | (6.670) |
| $\Delta Competition^{input}$ | -18.722** | -19.034*** | -17.716* | -23.331** |
| | (7.512) | (6.398) | (8.784) | (8.334) |
| $\Delta Competition^{output}$ * $Heavy$ | | | -1.274 | -2.864 |
| | | | (2.347) | (2.054) |
| $\Delta Competition^{input}$ * $Heavy$ | | | -10.731** | -5.260 |
| | | | (4.199) | (3.725) |
| $\Delta DemandConn$ * $Heavy$ | | | 12.035** | 8.424 |
| | | | (4.941) | (5.031) |
| $\Delta Competition^{output}$ * $TS$ | | | -3.170 | -5.087 |
| | | | (6.286) | (8.441) |
| $\Delta Competition^{input}$ * $TS$ | | | -8.447 | -17.166* |
| | | | (14.969) | (9.651) |
| $\Delta DemandConn$ * $TS$ | | | 4.962 | -23.892 |
| | | | (21.397) | (17.969) |
| $\Delta Competition^{output}$ * $Sh.Transp.$ | | | -94.062 | -51.113 |
| | | | (136.513) | (339.853) |
| $\Delta Competition^{input}$ * $Sh.Transp.$ | | | 1,365.574** | 1,069.165* |
| | | | (647.732) | (611.855) |
| $\Delta DemandConn$ * $Sh.Transp.$ | | | -1,313.302* | -1,091.214 |
| | | | (752.501) | (817.265) |
| Controls | NO | NO | YES | YES |
| Province FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Observations | 1,100 | 1,100 | 1,100 | 1,100 |
| Adj. R-squared | 0.323 | 0.269 | 0.354 | 0.261 |

*Notes:* "$\Delta CV$ emp" refers to the midpoint formula change in CV index measure of firm sizes by employment. "$\Delta CV$ sales" refers to the midpoint formula change in CV index measure of firm sizes by sales. Industry characteristics, *i.e.* average log weight-to-value ratio ($Heavy$), time sensitivity ($TS$), and share of expenditure on transportation ($Sh.Transp.$) are interacted with connectivity measures in columns (3) and (4). All observations are weighted by total employment in province-industry pairs. Robust standard errors are clustered at the province–industry level. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively.

Table B11: The effect of connectivity on size dispersion - dispersion calculated as Theil index, excluding Istanbul, including control variables.

| Dependent variable:<br>d size dispersion | (1)<br>$\Delta Thl$ emp | (2)<br>$\Delta Thl$ sales | (3)<br>$\Delta Thl$ emp | (4)<br>$\Delta Thl$ sales |
|---|---|---|---|---|
| $\Delta Competition^{output}$ | -11.703* | -11.629* | 4.294 | 4.976 |
| | (5.726) | (6.013) | (5.709) | (5.466) |
| $\Delta Competition^{input}$ | -21.906** | -21.930** | -16.963** | -17.981** |
| | (8.928) | (8.574) | (7.615) | (6.411) |
| $\Delta Competition^{output} * Heavy$ | | | -1.238 | -2.873 |
| | | | (2.789) | (2.584) |
| $\Delta Competition^{input} * Heavy$ | | | -5.597 | -1.056 |
| | | | (4.359) | (3.193) |
| $\Delta DemandConn * Heavy$ | | | 7.120 | 3.470 |
| | | | (5.668) | (5.439) |
| $\Delta Competition^{output} * TS$ | | | -7.968 | -5.393 |
| | | | (8.637) | (8.792) |
| $\Delta Competition^{input} * TS$ | | | 11.262 | 17.134 |
| | | | (16.334) | (17.801) |
| $\Delta DemandConn * TS$ | | | -10.412 | -21.318 |
| | | | (21.803) | (26.188) |
| $\Delta Competition^{output} * Sh.Transp.$ | | | -282.783 | -247.104 |
| | | | (256.722) | (311.592) |
| $\Delta Competition^{input} * Sh.Transp.$ | | | 1,055.810* | 580.207 |
| | | | (581.436) | (503.759) |
| $\Delta DemandConn * Sh.Transp.$ | | | -854.857 | -329.525 |
| | | | (687.758) | (651.442) |
| Controls | NO | NO | YES | YES |
| Province FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Observations | 1,100 | 1,100 | 1,100 | 1,100 |
| Adj. R-squared | 0.313 | 0.296 | 0.340 | 0.345 |

*Notes:* "$\Delta Thl$ emp" refers to the midpoint formula change in Theil index measure of firm sizes by employment. "$\Delta Thl$ sales" refers to the midpoint formula change in Theil index measure of firm sizes by sales. Industry characteristics, *i.e.* average log weight-to-value ratio ($Heavy$), time sensitivity ($TS$), and share of expenditure on transportation ($Sh.Transp.$) are interacted with connectivity measures in columns (3) and (4). All observations are weighted by total employment in province-industry pairs. Robust standard errors are clustered at the province–industry level. ***, **, * indicate the significance at 1%, 5%, 10% levels, respectively.