Price Discrimination, Backhaul Problems, and Trade Costs: Theory and Evidence from E-commerce Delivery*

Chunmian Ge[⋆] Hanwei Huang[†] Chang Liu[§] Wenji Xu[◇]

December 2024

Abstract

This paper studies the pricing of transportation services and the associated welfare implications. We develop a theoretical model to analyze two-part pricing by transportation service providers who face the backhaul problem and heterogeneous demands within and across regions. The model generates asymmetric transportation prices between regions and predicts that reductions in shipping costs lead to attenuated and unequal welfare gains. Using newly collected data from the largest platform for Chinese e-commerce delivery firms, we quantify the model and find that a universal 10% reduction in shipping costs leads the marginal delivery price to drop by 2.35%, the fixed fee to increase by 0.82%, delivery firms' profit to increase by 1.17%, and e-commerce sellers' welfare to rise by 0.34%, on average, with sellers in eastern and coastal cities gaining more than those in western and hinterland cities.

KEYWORDS: Trade costs, transportation, backhaul problem, price discrimination, two-part tariffs, welfare gains, e-commerce.

JEL CLASSIFICATION: F12, F14, L12, R12, R40

---

*For helpful comments, we would like to thank Treb Allen, Yan Bai, Ruoran Cheng, Kerem Coşar, Evan Cunningham, Jingting Fan, Brian Fujiy, Dennis Novy, Yong Wang, Chang Sun, Junji Xiao, Pai Xu, Vivian Yue, Ying Zhou, and the seminar participants at the City University of Hong Kong, IAAE 2024 and UEA 2024. We are responsible for any remaining errors. We gratefully acknowledge funding support from the National Natural Science Foundation of China (Grant 72272055), the Research Grant Council of Hong Kong (Project No. CityU 11503222), and the City University of Hong Kong (Project No. 7200695).

⋆ School of Business Administration, South China University of Technology. Email: bmgecm@scut.edu.cn.

† City University of Hong Kong, CEP, and CIFER. Email: huang.hanwei@cityu.edu.hk.

§ Department of Economics and Finance, City University of Hong Kong. Email: cliu542-c@my.cityu.edu.hk.

◇ Department of Economics and Finance, City University of Hong Kong. Email: wenjixu@cityu.edu.hk.

1 Introduction

Trade costs play a crucial role in shaping international and inter-regional trade and economic development. These costs encompass various factors, including transportation costs, policy barriers, contract enforcement costs, legal and regulatory costs, information costs, and local distribution costs (Anderson and Van Wincoop, 2004). Among these, transportation costs stand out as a particularly significant factor. With the rise of e-commerce, which largely bypasses tariffs (Fajgelbaum and Khandelwal, 2024) and other traditional trade barriers (Bai et al., 2020), efficient and affordable transportation of goods across borders and regions has become essential for firms to stay competitive and for consumers to access diverse products.

While numerous factors influence transportation costs, one critical friction unique to transportation is the backhaul problem. The backhaul problem arises when shipping flows are not balanced between two locations (Wilson, 1987): carriers who commit to the maximum capacity for a round trip inevitably have to run empty or partially loaded along routes with smaller flows, resulting in wasted resources and increased costs. This issue is particularly significant in international and inter-regional trade due to imbalances in trade flows. The backhaul problem both hampers the efficiency of transportation firms and adds to the overall trade costs, reducing aggregate welfare.

In addition to the backhaul problem, price discrimination considerations also play a crucial role in transportation service pricing. Due to differences in market size, income, and technologies, the demand for transportation services can exhibit significant heterogeneity across and within regions.[1] Given the heterogeneity in demand, transportation firms develop targeted pricing strategies and allocate their shipping capacities to meet the diverse needs of different regions and customer segments. To exploit demand heterogeneity *within* regions, transportation firms can engage in second-degree price discrimination by offering different prices based on service volume purchased. On the other hand, when faced with heterogeneous demands *across* regions, transportation firms can engage in third-degree price discrimination by offering different prices in different regions. Price discrimination practices also influence trade imbalances across regions and must, therefore, be considered alongside the backhaul problem in transportation firms' pricing strategies.

This paper provides the first comprehensive study of how the backhaul problem interacts with both second-degree and third-degree price discrimination considerations in affecting the pricing of transportation services, trade flows, and welfare.

We first build a novel theoretical model in which transportation firms face the backhaul problem, set two-part prices, and engage in third-degree price discrimination across different regions. The market size and composition vary across regions, which naturally gives rise to the backhaul

[1] For instance, in the e-commerce delivery market, urban areas with a higher concentration of tech-savvy consumers may exhibit a greater demand for inbound delivery services than areas with lower digital penetration.

problem: shipping flows from regions with higher demand for transportation services to regions with lower demand would exceed the return-trip flows. To exploit the demand heterogeneity within a given region across transportation service users, transportation firms implement second-degree price discrimination by offering *two-part prices*, which involves charging customers a fixed fee and a variable fee based on the quantity of goods delivered. To exploit the demand heterogeneity across regions, transportation firms can also implement third-degree price discrimination by offering different two-part prices in different regions. We characterize transportation firms' optimal pricing strategies and show that the backhaul problem and third-degree price discrimination can lead to asymmetric two-part prices and, hence, asymmetric trade costs between regions. We further decompose the asymmetric two-part prices into a component related to third-degree price discrimination and a component pertaining to the backhaul problem. Intuitively, heterogeneous demands across regions give rise to different pricing markups (price discrimination), while imbalances in shipping flows between regions result in asymmetric marginal shipping costs and, thus, asymmetric transportation prices (backhaul problems).

Our model implies that reductions in marginal shipping costs have *incomplete* pass-throughs to marginal transportation prices and tend to *increase* the fixed transportation prices. Both effects attenuate the welfare gains of the transportation service users. In particular, due to the backhaul problem, if the transportation flow along a shipping route falls short of the transportation firm's maximum shipping capacity, the marginal transportation price does not depend on the marginal shipping cost, leading to a zero pass-through from shipping costs to marginal prices. Even when the transportation flow along a shipping route reaches the firm's maximum shipping capacity, pass-throughs from marginal cost reductions to marginal prices are still incomplete due to markup pricing. Therefore, large reductions in the marginal shipping cost may lead to little or small changes in the marginal transportation prices. Furthermore, we show that the pass-through from the marginal shipping costs to the fixed fees is negative. As a result, transportation firms raise the fixed fee when the marginal cost declines. This allows transportation firms to extract more revenue, further limiting the welfare gains generated by reduced shipping costs.

To quantify the theoretical predictions, we focus on the Chinese e-commerce delivery market, which provides an appropriate setting for our empirical analysis. China is one of the largest and fastest-growing markets for e-commerce trade and delivery (Figure 1a).[2] However, the spatial distribution of e-commerce activities in China is highly uneven. As shown in Figure 1b, which depicts parcel throughput (total parcel inflows and outflows) across Chinese cities in 2019, coastal and eastern cities processed the majority of parcels.[3] The uneven spatial distribution of e-commerce

[2]China accounted for 45.8% of global online retail in 2017 (The World Bank and Alibaba Group, 2019) and 68.9% of global parcel volume in 2022 (Pitney Bowes, 2023).

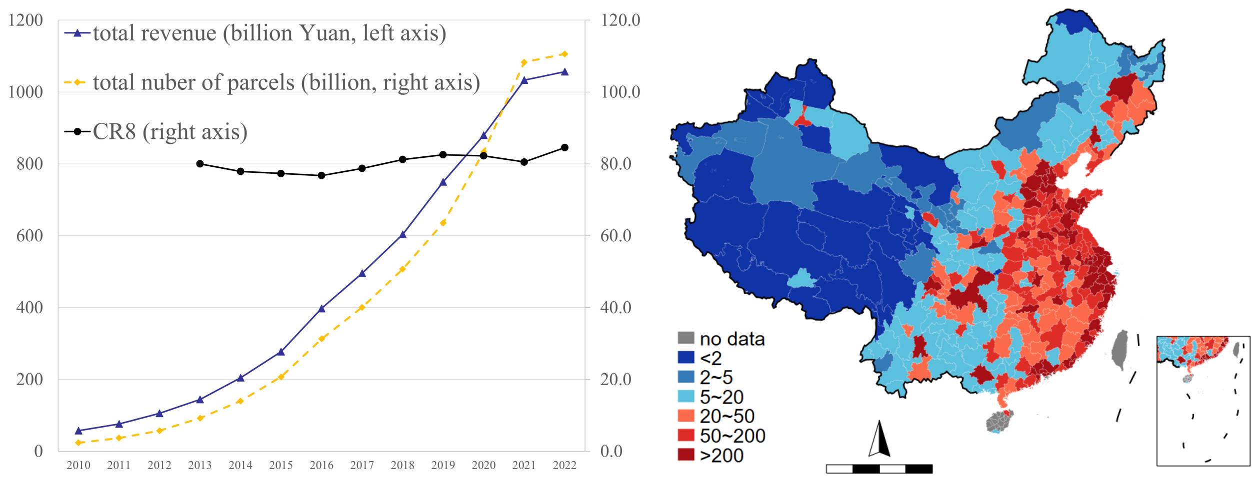
[3]Appendix Figure A.1 plots the spatial distribution of the number of Taobao towns (clusters of rural e-tailers) per city in China, which is also strikingly uneven across the country.

activities gives rise to imbalances in shipping flows between regions and exacerbates the backhaul problem faced by e-commerce delivery firms, as we provide direct empirical evidence in Section 2. In addition, the e-commerce delivery market in China is dominated by a few large firms. Despite the rapid growth of the e-commerce market, the top delivery companies have maintained stable market shares over the past decade. As Figure 1a indicates, the combined market share of the top 8 companies (CR8) remained at around 80% from 2013 to 2022. In this highly concentrated market, the leading delivery firms have considerable scope to exercise their market power and implement price discrimination both across and within Chinese cities.

Figure 1: E-commerce Delivery in China

(a) Market Size and Market Concentration

(b) Parcel Throughput per City



Notes: In Figure (a), CR8 measures the market share of top 8 firms. Figure (b) plots each city's parcel throughput (inflows plus outflows; unit: million) in 2019. Data sources: State Post Bureau and National Bureau of Statistics.

We use a novel and representative dataset from *Cainiao*, the largest platform for Chinese e-commerce delivery firms, to quantitatively evaluate our theoretical predictions. The platform facilitates trade between e-commerce sellers and delivery companies for bulk e-commerce shipping. We collected listed delivery prices among all mainland Chinese cities for February 2020 from this platform. Consistent with the two-part pricing assumption in our theoretical model, the observed prices have two components: a marginal price (e.g., Yuan/kilogram) and a fixed fee, specifying the minimum payment for each delivery. In addition to the posted prices, we also observe the number of delivery transactions across various routes.

Using the newly collected data, we document three stylized facts about e-commerce delivery that motivate our theoretical model and are used to validate the quantified model: (i) the volumes of e-commerce delivery among Chinese cities are imbalanced, with flows out of large cities exceeding flows out of small cities; (ii) the fixed and variable parts of listed e-commerce delivery prices in and out of cities and between city pairs are asymmetric; and (iii) listed delivery prices vary with differences in market size and the number of competing delivery firms.

Building on the characterization results of our theoretical model, we develop a new structural approach to estimate and validate key model parameters using our newly collected data. Specifically, we employ a block-recursive method to estimate e-commerce sellers' demand for delivery services in each city and the marginal shipping cost between city pairs, using the data on delivery prices and transactions. Our findings reveal that e-commerce sellers in coastal and eastern regions exhibit higher demand for delivery services compared to those in western and hinterland regions. We validate these estimation results by assessing the model's performance in explaining non-targeted moments. The quantitative model effectively captures the stylized facts discussed above. Moreover, the model's predictions for spatial distribution of e-commerce sellers in China and delivery firms' revenues within each city align closely with the observed data.

Using the estimated model, we conduct two quantitative exercises. First, we evaluate the roles of price discrimination and the backhaul problem in explaining asymmetric delivery prices. Our findings indicate that the asymmetry in marginal prices is primarily driven by the backhaul problem, while the asymmetry in fixed fees is largely attributed to third-degree price discrimination.

Finally, we conduct a counterfactual analysis to evaluate the market impact of shipping cost reduction. Our results show that under a 10% universal reduction in the marginal cost of shipping, the marginal delivery price decreases by an average of 2.35%, indicating that the pass-through from marginal shipping cost to marginal delivery price is far from complete. Additionally, the fixed fee *rises* by an average of 0.82% in response to the decline in marginal shipping costs. This suggests that delivery firms capture a significant portion of the e-commerce sellers' welfare gains from lower marginal prices through increased fixed fees. Across all city pairs, e-commerce sellers' surplus increases by an average of only 0.34%, with sellers in eastern and coastal cities benefiting more than those in western and hinterland regions. On average, delivery firms' revenues rise by 0.07%, while their profits increase by 1.17%.

This paper relates to several strands of research. First, our paper contributes to the literature on trade costs in international and inter-regional trade. It is well-known that trade costs between regions are asymmetric, which has far-reaching implications for income, welfare, and trade policies (Waugh, 2010; Bergstrand et al., 2013; Cuñat and Zymek, 2024). The existing literature typically attributes such asymmetric trade costs to trade imbalances and the backhaul problem (Behrens and Picard, 2011; Jonkeren et al., 2011; Ishikawa and Tarui, 2018; Wong, 2022). At the same time, it is recognized that firms in the transportation industry wield considerable market power, and variations in market competition across routes lead to significant differences in transportation service prices (Hummels et al., 2009; Asturias, 2020; Brancaccio et al., 2020; Allen et al., 2024b; Ignatenko, 2024). Our paper provides a unified framework that integrates both perspectives, enabling a quantitative assessment of the relative significance of the backhaul problem versus price discrimination in explaining asymmetric trade costs.

In addition, our findings contribute to the recent literature that examines the structure of trade costs. While the above-mentioned studies primarily focus on per-unit trade costs, there are mounting theoretical findings and empirical evidence highlighting the deviation from the iceberg trade cost assumption (Irarrazabal et al., 2015; Bosker and Buringh, 2020) and importance of fixed trade costs (Melitz, 2003; Antràs et al., 2017). The existing literature often attributes fixed trade costs to factors such as marketing expenses, distribution channel setup, adaptation to policy regulations, and information frictions (Roberts and Tybout, 1997; Melitz, 2003; Arkolakis, 2010; Allen, 2014; Huang et al., 2024). We find that transportation costs alone have a fixed component, as a result of the price discrimination practice by transportation firms. Additionally, our results indicate that as the per-unit transportation cost decreases, the fixed component of transportation cost increases, partially offsetting the welfare gains for market participants from the reduction in per-unit trade costs. This effect is quantitatively significant in our estimated model.

More broadly, our research relates to the literature examining various other factors that influence trade costs. A large body of work has examined the role of infrastructure such as ports, highways, and railways (Limao and Venables, 2001; Duranton et al., 2014; Faber, 2014; Coşar and Demir, 2016; Ghani et al., 2016; Donaldson, 2018; Fan et al., 2023; Ma and Tang, 2024), technologies such as containerization (Levinson, 2016; Coşar and Demir, 2018), and the configuration of transportation networks (Fajgelbaum and Schaal, 2020; Allen and Arkolakis, 2022; Alder et al., 2023) in affecting trade costs. Our results show that the pass-through of infrastructure improvements to transportation costs is incomplete and uneven across regions, which depends on the market power of transportation service providers and the nature of trade imbalances.

Our paper also contributes to the literature on the backhaul problem, which arises when trade flows are imbalanced between two regions. This problem has been examined in the airline industry (Evans and Kessides, 1994), maritime shipping Wong (2022), and trucking (Harris and Nguyen, 2024). Other than posing a serious challenge to the transportation industry by complicating transportation routing and planning (Cordeau et al., 1998; Holmberg et al., 1998), the backhaul problem has important implications for trade costs and economic geography (Behrens and Picard, 2011; Behrens et al., 2018), trade and industrial policies (Ishikawa and Tarui, 2018; Friedt and Wilson, 2020), and trade patterns and welfare (Brancaccio et al., 2020; Cuñat and Zymek, 2024). Our contribution is to first show that the backhaul problem leads to asymmetry in not only the per-unit shipping price, as shown in Behrens and Picard (2011) and Ishikawa and Tarui (2018), but also in fixed fees. We further show that the backhaul problem leads to zero pass-through from marginal shipping costs to shipping prices, resulting in limited welfare gains when there is a reduction in shipping costs. Therefore, policies that alleviate the backhaul problem could further enhance the benefits of improved transportation infrastructure.

In terms of our empirical setting, this paper also contributes to the literature that evaluates the welfare gains from the expansion of e-commerce. For instance, Fan et al. (2018) suggest that small cities in China gain more from e-commerce than large cities. Couture et al. (2021) find that younger and richer Chinese households in remote markets gain significantly from e-commerce expansion by overcoming logistical barriers to e-commerce. Dolfen et al. (2023) examine U.S. credit and debit card data from Visa and find that consumers in more densely populated counties gain more from e-commerce. We show how unequal gains from e-commerce within and across regions can arise due to the backhaul problem and price discrimination in e-commerce delivery. For example, as sending parcels to smaller markets is more expensive than to larger markets due to the backhaul problem, e-commerce consumers in smaller markets benefit less from reductions in the marginal cost of shipping goods across cities.

The rest of the paper proceeds as follows. Section 2 discusses the research background, the data sources, and three motivating facts. Section 3 introduces the theoretical model and characterizes the optimal pricing of transportation services. Section 4 estimates the model and validates the estimation results. Section 5 quantifies the importance of price discrimination and the backhaul problem in explaining the observed price asymmetry and conducts counterfactual analyses on the market impacts of reductions in shipping costs. Section 6 concludes the paper.

2 Background, Data, and Motivating Evidence

In this section, we first introduce the background of our empirical setting, i.e., e-commerce and the e-commerce delivery industry in China, and discuss the data used in this paper. We then present some stylized facts that motivate our theoretical and quantitative analyses.

2.1 Background and Data

E-commerce and the e-commerce delivery industry have proliferated over the past decade and become important sectors in the Chinese economy. Between 2013 to 2023, online retail sales in China grew at a remarkable compound annual rate of 23%, soaring from 1.8 trillion Chinese Yuan (CNY) to 15.4 trillion CNY. By 2023, 27.6% of the country's total retail sales of consumer goods occurred online, up from a mere 1.1% in 2008. That year, 84.4% of online retail sales were for physical products that required shipping, leading to the delivery of 132.1 billion parcels. Notably, 87.4% of these parcels were inter-city deliveries, which are the focus of our study.[4]

[4]Of the remaining parcels, 10.3% were intra-city, and 2.3% were for overseas destinations. Source: National Bureau of Statistics and State Post Bureau of China.

According to Fitch Ratings (2020), e-commerce delivery in China is dominated by a few large domestic private logistics companies, which command a large market share (see Figure 1a). The state-owned China Post mainly serves government and public sector needs, while foreign companies like DHL focus on international deliveries. Parcels are primarily transported overland by trucks.[5] These delivery companies typically own or rent tens of thousands of trucks running between cities. Hence, it is quite different from the one-man and one-truck shipping industry in Colombia studied by Allen et al. (2024b).

Our primary data source is *Cainiao*, an online platform for delivery companies to provide bulk e-commerce shipping services. Cainiao was founded by the *Alibaba* Group, one of the largest players in the Chinese e-commerce market. To facilitate e-commerce trade, Cainiao provides a platform for delivery companies to list prices and shipping services across all Chinese cities. E-commerce sellers, most likely from B2B platform, *1688.com*, and B2C platforms, *Taobao* and *Tmall*,[6] are then able to find delivery companies on the platform and use their delivery services. For large orders, they can choose full-truckload services ordering a whole truck, for smaller ones, they can use less-than-load services. For small parcels (typically $\leq$ 30 kilograms), they can use express services. In 2017, Cainiao and its network of delivery firms facilitated around 70% of China's e-commerce deliveries (Fitch Ratings, 2020).[7]

We collect the price listings from delivery companies for each shipping route for all mainland Chinese cities from Cainiao in February 2020.[8] As can be seen from panel (a) of Table 1, which shows the summary statistics at the route level (city to city), the price listings take the form of two-part prices, consisting of a fixed fee and a constant marginal price. In total, we have a sample of 81 companies and their listings along routes connecting 334 Chinese cities, which covers the universe of firms and price listings on the platform. Consistent with the high concentration of the market, we find that the average number of firms per route is about 9 and the median is 8. From the platform, we also observe the number of delivery transactions on each route in recent months. Finally, we use the Baidu Map API service to capture the driving distance between all pairs of Chinese cities as a proxy for the transportation costs between the city pairs.

To complement the analysis, we also collect several datasets to capture the various character-

[5] Only two shipping companies, SF Express and YTO Express, operate air shipping fleets; however, even for SF Express, the leader in air deliveries, air shipments accounted for only 22% of its total deliveries in 2018, with railway shipping accounted for just 1.2%. As of 2022, 73.3% of goods in China were transported via road.

[6] 1688.com, Taobao, and Tmall are all owned by Alibaba, are among the largest Chinese online retail and wholesale platforms. Their combined market share is around 57% in 2020.

[7] Off-platform e-commerce delivery services include deliveries offered by the second-largest e-commerce platform, JD.com, and other firms not in partnership with Alibaba. Social commerce platforms such as Douyin (TikTok's counterpart in China), which integrate social media with online retailing, have gained popularity in recent years. Nevertheless, they still rely on e-commerce delivery firms for parcel deliveries.

[8] We exclude listings for shipping to Hong Kong, Macau, and Taiwan, where e-commerce goods flows are subject to customs checks crossing the mainland border.

istics of different Chinese cities, as shown in panel (b) of Table 1. We retrieve the data on parcel throughput and revenue from parcel deliveries for each city from the annual report of the State Post Bureau of China. Since our sample is from February 2020, after the outbreak of the COVID-19 pandemic, we control for the presence of COVID-19 containment policies in Chinese cities using data from Ge et al. (2023).

Table 1: Summary Statistics of Route Level and City Level Data

	N	Mean	Std.	Median	Min	Max
Panel (a): route-level summary statistics						
marginal price	111220	2.21	0.93	2.03	0.56	8.37
fixed fee	111220	287.5	204.9	205.6	20	1795.4
number of firms	111220	9.13	4.39	8	1	32
ln(driving distance)	111220	7.33	0.68	7.41	2.98	8.72
delivery transactions	111220	3.69	156.13	0	0	47316
Panel (b): city-level summary statistics						
parcel throughput	334	18995.6	62131.0	2667.1	6.68	634680.3
parcel revenue	334	224081.6	911019.1	33698.5	189.1	12888000
presence of COVID-19 policies	334	0.40	0.49	0	0	1

Notes: Panel (a) presents summary statistics for variables at the route level (city to city). The marginal price (unit: Yuan/kilogram) and fixed fee (unit: Yuan) are the average values across different services and delivery firms operating on the same route from Cainiao. The number of delivery firms indicates the total number of firms operating within that route. The variable ln(driving distance) represents the logarithm of the driving distance between two cities on the route. Panel (b) provides summary statistics for variables at the city level for the year 2019. Parcel throughput refers to the total number of parcel inputs and outputs within a city. Parcel revenue represents the total income generated from delivery services within a city.

2.2 Stylized Facts

Fact 1 (E-commerce Delivery Imbalances): *E-commerce delivery transactions between cities are imbalanced, with flows out of large cities exceeding flows out of small cities.*

As discussed in the Introduction and as presented in Figure 1b, the spatial distribution of e-commerce activities in China is extremely uneven. In addition, the Chinese logistics network exhibits a *hub-and-spoke* structure, primarily centered around a few national core hubs and regional provincial trade hubs (Head et al., 2017; Li et al., 2023). Shipping between these hubs and regular nodes in the network is imbalanced, leading to substantial backhaul problems.

We present evidence of such imbalances in e-commerce trade using the delivery transaction data from Cainiao. As Panel (a) of Table 2 shows, on average, the absolute difference of delivery transactions in the two directions of shipping between a pair of cities is as large as 0.33 log points across all city pairs. Since Cainiao only reports the cumulative transaction volumes for the most recent few months, a significant fraction of the observed transaction volumes are zeros. However, the asymmetries in delivery transactions are not due to these zeros. For the 3252 pairs of cities where we observe positive numbers of delivery transactions in both shipping directions, the difference is even larger at 1.44 log points. Thus, imbalances in shipping volumes are prevalent in e-commerce deliveries in China.

Table 2: Asymmetries in E-commerce Deliveries and Pricing

	N	Mean	Std.	Median	Min	Max
Panel (a): Absolute differences in delivery transactions between a pair of cities						
All city pairs	55609	0.33	0.76	0	0	9.60
City pairs with positive transactions both ways	3252	1.44	1.17	1.18	0	7.00
Panel (b): Absolute differences in delivery prices in and out of a city						
Marginal price	334	0.12	0.094	0.11	0.000095	0.65
Fixed fee	334	0.43	0.36	0.36	0.00032	2.00
Panel (c): Absolute differences in delivery prices between a pair of cities						
Marginal price	55609	0.18	0.15	0.13	0	1.36
Fixed fee	55609	0.58	0.62	0.31	0	3.80

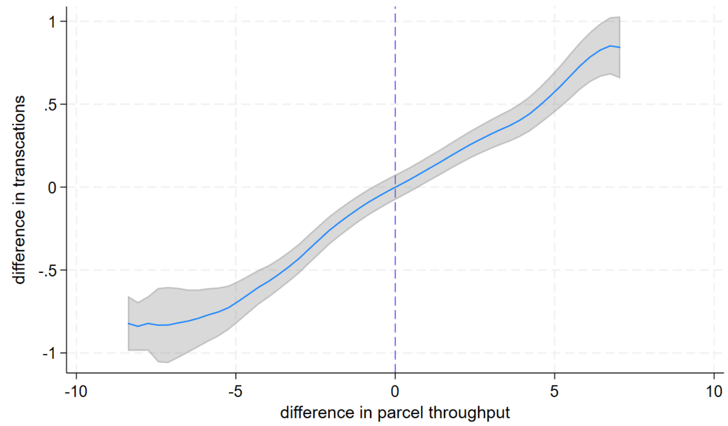
Notes: Panel (a) presents the magnitude of differences in delivery transactions (in logarithm) in the two shipping directions between city pairs. Panel (b) provides summary statistics for the magnitude of price differences (in logarithm) for deliveries in and out of each city. Panel (c) provides summary statistics for price differences (in logarithm) for deliveries between each pair of cities.

Furthermore, as indicated in Figure 2, we find a positive correlation between the difference in bilateral delivery transactions and the difference in delivery market sizes between the paired cities. Here, the market size of a city is measured by its parcel throughput, which is the total of all its inbound and outbound shipments. The observed delivery transactions are roughly balanced only between cities of similar size. Hence, given the hub-and-spoke transportation network, when delivery firms ship parcels from a large city to a small city, the return trip is likely to be partially loaded, resulting in underutilized transportation resources.

Fact 2 (Asymmetric Delivery Pricing): *The fixed and variable parts of listed prices for e-commerce deliveries in and out of cities, and between cities are asymmetric.*

Next, we consider the listed prices for e-commerce delivery. As mentioned above, we observe that delivery companies offer two-part prices with both fixed and variable components. We calculate the absolute differences in the prices for shipping in and out of each city (in logarithm) and present their distributional statistics in Panel (b) of Table 2. On average, the listed marginal price (i.e., the variable component of delivery price) differs by 0.12 log points for shipping in and out of Chinese cities. The asymmetry in the fixed component, which we refer to as the fixed fee, is even greater, with an average of 0.43 log points and a maximum of 2 log points. Figure 3a shows the kernel density of absolute price differences, which reveals a wider dispersion in fixed fee differences compared to the dispersion in marginal price differences.

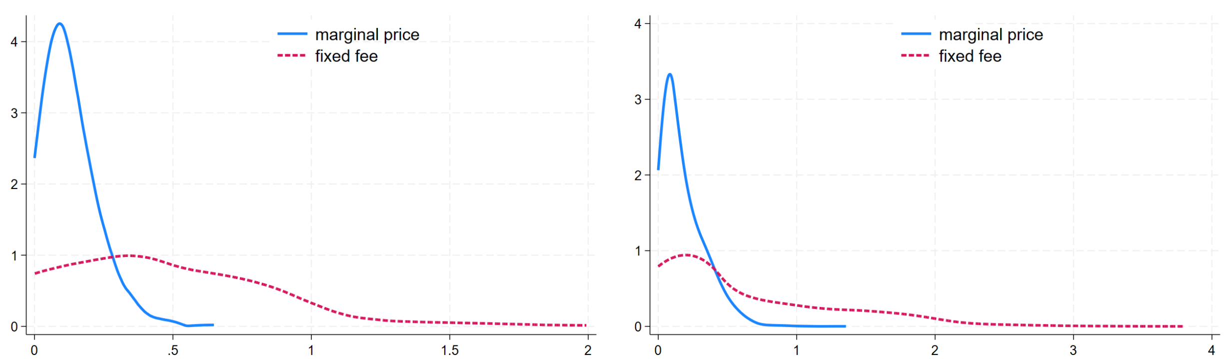
Figure 2: E-commerce Delivery Imbalances and Market Sizes



Notes: This figure plots the local polynomial regression of the difference in the number of bilateral delivery transactions on the difference in parcel throughputs between city pairs. The horizontal axis is the difference in the parcel throughput. The vertical axis is the difference in the number of delivery transactions between city pairs. The solid line is the best-fit line. The shaded areas are the 95% confidence intervals.

Figure 3: Kernel Density of Absolute Price Differences

(a) city level (b) city pair level



Notes: (a) plots the kernel density of absolute differences in delivery prices in and out of each city. (b) plots the kernel density of the absolute differences in the delivery prices in the two directions of deliveries between each pair of cities.

As shown in Panel (c) of Table 2, the asymmetry in listed delivery prices is even more pronounced at the city-pair level. The average absolute difference in the marginal prices for the two shipping directions between a pair of cities is 0.18 log points. Again, the asymmetry in fixed fees is more significant than that in marginal prices. The average absolute difference in the fixed fees for the two shipping directions between a city pair is 0.58 log points. Figure 3b plots the kernel density of these differences, which confirms that the differences in fixed fees are more dispersed than those in marginal prices.

Fact 3 (Price Variation): *Listed delivery prices vary with differences in market sizes and the number of competing delivery firms.*

We next examine how listed delivery prices vary with city and shipping route characteristics, which provides suggestive evidence of how the consideration of the backhaul problem and market power affects delivery firms' pricing strategies. Formally, we study the determinants of price differences (in logarithm) between the two shipping directions of a city pair by conducting the following regression:

$$\ln \frac{p_{ij}}{p_{ji}} = \eta \ln \frac{MS_i}{MS_j} + \beta(X_{ij} - X_{ji}) + \gamma Z_{ij} + \xi_{ij}. \tag{1}$$

where p_{ij} is the price for shipping parcels from city i to city j. As suggested by Figure 2, differences in market sizes generate imbalances in delivery flows, leading to the backhaul problem. The first term on the right-hand side of equation (1), $\ln(MS_i/MS_j)$, represents the difference in market size between cities i and j, which helps capture the impact of the backhaul problem on delivery prices. The second term, $X_{ij} - X_{ji}$, represents the difference in the number of delivery firms serving each shipping route between the city pair, which helps capture differences in the firms' market power in the two shipping directions.[9] The third term, Z_{ij}, is a dummy that controls for the presence of COVID-19 control policies in either city, as our sample period is after the pandemic outbreak. Finally, ξ_{ij} is a statistical error term.

Table 3: Delivery Firms' Two-Part Tariff Pricing

	difference in marginal price			difference in fixed fee		
	(1)	(2)	(3)	(4)	(5)	(6)
difference in parcel throughput	0.0475^a	0.0524^a	0.0524^a	0.155^a	0.0553^a	0.0553^a
	(0.0003)	(0.0005)	(0.0005)	(0.0012)	(0.0017)	(0.0017)
difference in firm #		-0.0441^a	-0.0440^a		0.902^a	0.902^a
		(0.0034)	(0.0034)		(0.0117)	(0.0117)
COVID-19 policy control	No	No	Yes	No	No	Yes
R-squared	0.333	0.335	0.335	0.268	0.346	0.346
No. of observations	55609	55609	55609	55609	55609	55609

Notes: This table presents estimation results of equation (1) using city-pair level data. The dependent variables are differences in delivery prices in the two directions of shipping between paired cities. Columns (1) - (3) examine the marginal delivery price, and columns (4)-(6) the fixed fee. Both the dependent and independent variables are in logarithm. The COVID-19 policy controls are dummies indicating the presence of COVID-19 policies in each city. The numbers in the parentheses are robust standard errors. a indicates a significance level of 0.01.

Table 3 presents the regression results where the outcome variables are the log difference in

[9]This measure is not zero because not all firms serve both ways of a shipping route. In the data, about one third of the time, firms serve only one way of the trade routes between paired cities.

marginal prices in columns (1) - (3) and the log difference in fixed fees in columns (4) - (6), respectively. We find that the regression coefficient of $\ln(MS_i/MS_j)$ is positive and significantly different from zero. The preferred specifications are columns (3) and (6), in which we added all the controls. According to the results, a 1% increase in the difference between market sizes of the origin and destination cities is associated with about a 5.2% increase in the difference in marginal prices and a 5.5% increase in the difference in fixed fees. Therefore, the backhaul problem is likely an important concern in delivery pricing. Delivery firms tend to charge higher prices for shipping from a large market to a small market than from a small market to a large market, both in terms of the marginal price and fixed fee.

As shown in Table 1, there is considerable variation in the number of delivery firms serving different shipping routes. We exploit this variation in market competitiveness to examine how delivery firms' market power affects delivery pricing. As shown in Table 3, the regression coefficient for the difference in the log number of delivery firms is negative for the marginal price in columns (2) and (3) and positive for the fixed fee in columns (5) and (6). Thus, delivery firms tend to charge a higher marginal price and a lower fixed fee when they have more market power.

Robustness Checks We have got access to post-COVID-19 data from Cainiao for May 2024. Nevertheless, Cainiao has stopped sharing the number of e-commerce delivery transactions in the public domain. Hence, we prefer the data from the early COVID-19 period. Nevertheless, our findings regarding pricing in Fact 2 and Fact 3 remain robust using the post-COVID-19 data, as shown in Appendix Figure A.2 and Table A.1.

3 Theoretical Model

The previous section highlights the need for a deeper and renewed examination of optimal pricing by transportation firms that face trade imbalances and the backhaul problem. Our data show that e-commerce delivery firms offer two-part prices, with both fixed and variable components that feature systematic asymmetries influenced by the backhaul problem and market competitiveness. Nevertheless, the existing literature on transportation service pricing and its impact on trade costs typically assumes that transportation firms adopt linear pricing policies, even though freight rates are usually nonlinear in trade volume. (Slack and Gouvernal, 2011; Ignatenko, 2024).

To formally examine the two-part pricing of transportation services under trade imbalances, we next introduce a model of the delivery market wherein delivery service providers face imbalances in shipping demands and adopt two-part tariffs to price their services in the presence of backhaul problems and heterogeneous consumer demands in different regions. The model generates novel predictions about how pricing patterns and welfare gains vary across regions and how the fixed and variable components of the delivery price react *differently* to changes in shipping costs. These theoretical predictions elucidate the economic forces driving the pricing patterns documented in the previous section and form the foundation of the quantitative analysis in the next section. To align with our empirical context, the model will refer to different geographical regions as "cities" and buyers of delivery services as "e-commerce sellers."

3.1 The Setting

Demand for Delivery Service There is a finite set of cities, denoted by $\mathcal{C}$. In each city $i \in \mathcal{C}$, a mass $N_i \geq 0$ of e-commerce sellers demand delivery services to ship parcels to other cities. Each e-commerce seller's demand depends on a "willingness-to-pay" parameter θ, which is independent and identically distributed across the e-commerce sellers within the same city according to the Pareto distribution on $[\underline{\theta}_i, \infty)$ with probability density function

$$f_i(\theta) = \frac{\alpha_i \underline{\theta}_i^{\alpha_i}}{\theta^{\alpha_i+1}}.^{10} \tag{2}$$

We refer to θ as the e-commerce seller's "type" and assume $\alpha_i > 2$, so the type distribution has a finite expectation and a finite variance.

Without loss of generality, we assume that each e-commerce seller submits one bulk delivery order between each pair of two cities. Following Tirole (1988), we assume that a type-θ e-commerce seller in city i derives the following utility from delivering parcels to city j

$$U_{ij}(\theta, q) = \frac{\theta^2 - (\theta - q)^2}{2} - T_{ij}(q), \tag{3}$$

where q is the quantity of delivery service demanded (which can consist of multiple parcels in our empirical setting) and $T_{ij}(q)$ is the payment for the delivery service. The e-commerce sellers choose quantity q to maximize their utilities, given the payment policy $T_{ij}(\cdot)$ imposed by the delivery service provider, which we introduce next.

Supply of Delivery Service There are $n_{ij} \in \mathbb{N}_+$ symmetric delivery firms that provide shipping services for e-commerce sellers between each city pair $(i, j) \in \mathcal{C}^2$. Since e-commerce sellers often maintain long-term relationships with delivery firms due to the benefits of stability and efficiency, we focus on the symmetric equilibrium such that delivery firms have equal market shares, each facing a residual demand with a Pareto distribution given by (2) (see Section A.5 for a microfoun-

[10] Such power-law distributions are widely used (see Gabaix (2016)). According to Huang et al. (2021), the sales distribution of Chinese e-commerce sellers closely follows a Pareto distribution.

dation of this assumption).[11] The analysis below therefore focuses on the pricing problem of an individual delivery firm.

Suppose that each firm has a constant marginal cost c_{ij} of expanding its shipping capacity between cities i and j. Due to the backhaul problem, the firm has to maintain a shipping capacity to meet the maximum delivery demand between cities i and j.[12] In addition, the firm faces a fixed cost of f_{ij}, which is not related to the number of deliveries (e.g., tolls). Thus, the total cost to an individual firm to provide delivery services between cities i and j is

$$c_{ij} \cdot \frac{1}{n_{ij}} \cdot \max(q_{ij}, q_{ji}) + f_{ij}, \tag{4}$$

where q_{ij} and q_{ji} are the aggregate quantities of e-commerce shipping from city i to city j and from city j to city i, respectively.[13] We assume $0 < c_{ij} < \min\{\frac{\alpha_i - 2}{\alpha_i - 1}\underline{\theta}_i, \frac{\alpha_j - 2}{\alpha_j - 1}\underline{\theta}_j\}$ to ensure that the firm's optimal pricing strategy is interior.

The firm adopts a two-part tariff pricing strategy for its delivery services. Specifically, the firm charges a fixed fee F_{ij} (or F_{ji}, respectively) for submitting an order and a marginal price p_{ij} (or p_{ji}, respectively) for delivering each parcel from city i to city j (or from city j to city i, respectively). Hence, if an e-commerce seller in city i ships q units of goods to city j, the seller is charged the following payment by the delivery firm:

$$T_{ij}(q) = p_{ij}q + \mathbb{1}_{\{q>0\}}F_{ij}, \tag{5}$$

where $\mathbb{1}_{\{q>0\}}$ is an indicator that equals 1 if $q > 0$ and equals 0 otherwise. Similarly, in the opposite direction, if an e-commerce seller in city j ships q units of goods to city i, the seller pays $T_{ji}(q) = p_{ji}q + \mathbb{1}_{\{q>0\}}F_{ji}$ to the delivery firm.

The firm chooses p_{ij}, p_{ji}, F_{ij} and F_{ji} to maximize its total profit from serving e-commerce sellers in both city i and city j to ship their parcels to the other city.

[11] In Section A.5, we show that the same Pareto-shaped demand function in (2) emerges endogenously as the residual demand faced by each delivery firm in an oligopoly model where e-commerce sellers select and form long-term relationships with delivery firms. It is innocuous to assume that delivery firms have equal market shares because asymmetric market shares can also arise in equilibrium but lead to the same pricing and welfare outcomes and are merely artifacts of e-commerce sellers' tie-breaking rules. We assume that the delivery firms have market power because, in a transportation industry with backhaul problems, perfect competition would lead to a zero marginal price for one shipping direction (see, e.g., Behrens and Picard, 2011), which is inconsistent with our data.

[12] As discussed in Section 2.2, China's logistics network follows a hub-and-spoke configuration (Head et al., 2017; Li et al., 2023), and interregional trade flows are imbalanced. Therefore, the backhaul problem is likely significant in China's e-commerce delivery market. Allen et al. (2024a) studies a traveling trucker problem in which a trucker's itinerary may involve more than two cities. We are unable to incorporate multi-stop deliveries into our analysis due to the lack of data on delivery firms' itineraries.

[13] We follow the literature and assume that the marginal cost is symmetric between any two cities, i.e., $c_{ij} = c_{ji}$, because the forehaul and the backhaul between two cities should have similar driving distances.

3.2 Equilibrium and Optimal Pricing

Given the delivery firm's price offerings, e-commerce sellers determine how much delivery service to order, if any. According to (3) and (5), if a type-θ e-commerce seller in city i uses the delivery service to ship goods to city j, she will optimally choose to deliver a quantity

$$q = \max(\theta - p_{ij}, 0) \tag{6}$$

of parcels to maximize her utility $U_{ij}(\theta, q) - p_{ij}q - \mathbb{1}_{\{q>0\}}F_{ij}$. E-commerce sellers with types below p_{ij} will not use the delivery service even if the fixed fee is zero. By paying the fixed fee F_{ij} and optimally choosing a quantity $\theta - p_{ij}$ of delivery service, an e-commerce seller with type $\theta \geq p_{ij}$ obtains a surplus of

$$S(\theta, p_{ij}, F_{ij}) := \max_q [U_{ij}(\theta, q) - p_{ij}q - \mathbb{1}_{\{q>0\}}F_{ij}] = \frac{(\theta - p_{ij})^2}{2} - F_{ij}.$$

Let $\theta^*_{ij} \geq p_{ij}$ denote the seller type that is indifferent to using the delivery service or not, i.e., $S(\theta^*_{ij}, p_{ij}, F_{ij}) = 0$.[14] Hence, a type-$\theta$ e-commerce seller in city i pays the fixed fee F_{ij} and delivers goods to city j if and only if $\theta \geq \theta^*_{ij}$. E-commerce sellers with higher types demand higher quantities of delivery service and enjoy larger surpluses.

In aggregate, e-commerce sellers in city i demand the following quantity of delivery service to ship goods to city j:

$$q_{ij} = N_i \int_{\theta^*_{ij}}^{\infty} (\theta' - p_{ij}) f_i(\theta') d\theta' = \left(\frac{\alpha_i}{\alpha_i - 1} \frac{\underline{\theta}_i^{\alpha_i}}{\theta_{ij}^{*(\alpha_i - 1)}} - p_{ij} \left(\frac{\underline{\theta}_i}{\theta^*_{ij}} \right)^{\alpha_i} \right) N_i. \tag{7}$$

Given the e-commerce sellers' delivery decisions (as responses to the prices), the delivery firm chooses the pricing strategy $(p_{ij}, p_{ji}, F_{ij}, F_{ji})$ to maximize its profit:

$$\Pi_{ij} := \frac{1}{n_{ij}} \left[p_{ij}q_{ij} + p_{ji}q_{ji} + F_{ij}N_i \int_{\theta^*_{ij}}^{\infty} f_i(\theta')d\theta' + F_{ji}N_j \int_{\theta^*_{ji}}^{\infty} f_j(\theta')d\theta' - c_{ij} \max(q_{ij}, q_{ji}) \right] - f_{ij}. \tag{8}$$

To simplify the exposition, we impose the following assumption, which requires that the demands for delivery services are sufficiently imbalanced and the backhaul problem is salient.

Assumption 1. *The demand (for delivery service) between any pair of cities $(i, j) \in \mathcal{C}^2$ is such that either $N_i\underline{\theta}_i - N_j\underline{\theta}_j > N_i c_{ij}$ or $N_j\underline{\theta}_j - N_i\underline{\theta}_i > N_j c_{ij}$.*

This assumption is innocuous for both the theoretical and the quantitative analysis. When

[14] To simplify the exposition, we suppress the dependence of θ^*_{ij} on p_{ij} and F_{ij}.

this assumption is violated, it can be shown (theoretically) that the delivery firm always induces perfectly symmetric delivery volumes, i.e., $q_{ij} = q_{ji}$ (see Appendix A.1). As we want to focus on situations where the backhaul problem is important and all our data feature asymmetric service volumes, we will maintain Assumption 1 in our analysis.

The next result characterizes the delivery firms' equilibrium pricing strategies.

Proposition 1. *For any city pair $(i, j) \in \mathcal{C}^2$, in equilibrium, $q_{ij} \neq q_{ji}$, the marginal price is*

$$p_{ij} = \begin{cases} \frac{\underline{\theta}_i}{\alpha_i - 1} + c_{ij}, & \text{if } q_{ij} > q_{ji}, \\ \frac{\underline{\theta}_i}{\alpha_i - 1}, & \text{if } q_{ij} < q_{ji}, \end{cases} \tag{9}$$

and the fixed fee is

$$F_{ij} = \frac{(\underline{\theta}_i - p_{ij})^2}{2}. \tag{10}$$

Proof. *See Appendix A.1.*

The term $\frac{\underline{\theta}_i}{\alpha_i - 1}$ on the right-hand side of equation (9) captures the markup charged by the delivery firm to e-commerce sellers due to the firm's market power and reflects the effects of (third-degree) price discrimination. In particular, the delivery firm charges a higher markup if there is a larger density of e-commerce sellers with high willingness-to-pay, which happens when the Pareto distribution has a fatter right-tail (i.e., a smaller α_i) or features larger scales (i.e., a larger $\underline{\theta}_i$). The second term, the marginal shipping cost c_{ij}, only shows up for the shipping direction with a larger delivery volume, which reflects the impact of the backhaul problem on the marginal prices. When the service volumes between the cities are imbalanced ($q_{ij} \neq q_{ji}$), the backhaul problem leads to an increase in the marginal price only for the shipping direction with a larger delivery volume.

Similarly, from equation (10), we note that the delivery firm tends to charge a higher fixed fee on e-commerce sellers if the Pareto distribution has a larger scale (i.e., a larger $\underline{\theta}_i$). For the haul with a lower marginal price, the delivery firm extracts more surplus by setting a higher fixed fee.

The asymmetry in shipping prices between any two cities can then be attributed to the effects of either price discrimination or the backhaul problem, as demonstrated by the following result.

Corollary 1. *The difference in the marginal prices and the fixed fees for the two shipping directions between any two cities i and j can be decomposed into a component driven by price discrimination and another component driven by the backhaul problem:*

$$p_{ij} - p_{ji} = \underbrace{\frac{1}{\alpha_i - 1}\underline{\theta}_i - \frac{1}{\alpha_j - 1}\underline{\theta}_j}_{price\ \ discrimination} + \underbrace{(\mathbb{1}_{\{q_{ij} > q_{ji}\}} - \mathbb{1}_{\{q_{ij} < q_{ji}\}})c_{ij}}_{backhaul\ \ problem}, \tag{11}$$

and

$$F_{ij} - F_{ji} = \frac{(\underline{\theta}_i - p_{ij} + \underline{\theta}_j - p_{ji})}{2} \left[\underbrace{\frac{\alpha_i - 2}{\alpha_i - 1}\underline{\theta}_i - \frac{\alpha_j - 2}{\alpha_j - 1}\underline{\theta}_j}_{price\ \ discrimination} - \underbrace{(\mathbb{1}_{\{q_{ij} > q_{ji}\}} - \mathbb{1}_{\{q_{ij} < q_{ji}\}})c_{ij}}_{backhaul\ \ problem} \right], \quad (12)$$

where $\mathbb{1}_{\{q_{ij} > q_{ji}\}}$ is an indicator that equals 1 if $q_{ij} > q_{ji}$ and 0 otherwise.
Proof. *See Appendix A.2.*

Equation (11) decomposes the difference in the marginal prices between the two shipping directions into two components. The first component reflects the effect of price discrimination and captures the difference in markups resulting from the varying demands for delivery services by e-commerce sellers in the two cities. The delivery firm tends to charge a higher price in the city in which the e-commerce sellers have a higher average willingness to pay for delivery services. The second component captures the price difference caused by the backhaul problem when trade flows are imbalanced across the two shipping directions. The delivery firm tends to charge a higher marginal price for the haul with a greater overall trade flow.

Similarly, equation (12) provides the decomposition of the difference in the fixed fees. The term, $\frac{\underline{\theta}_i - p_{ij} + \underline{\theta}_j - p_{ji}}{2}$, is always positive (see Appendix A.2). Therefore, the difference in the fixed fees, $F_{ij} - F_{ji}$, can also be attributed to price discrimination and the backhaul problem, as reflected in the terms in the square brackets of equation (12). Due to price discrimination, the delivery firm tends to charge a higher fixed fee in the city with a higher willingness to pay. Due to the backhaul problem, the firm tends to lower the fixed fee for the haul with a greater overall trade flow.

The marginal prices and the fixed fees constitute the e-commerce trade costs between the cities, both of which depend on c_{ij}, the marginal shipping cost. We are interested in how a reduction in the shipping cost (e.g., due to improved transportation infrastructure) would translate into changes in the trade costs in terms of both the marginal price and the fixed fee for delivery services.

Corollary 2. *The pass-through rates of the marginal shipping cost c_{ij} to the marginal price and the fixed fee are*

$$\frac{\partial \ln p_{ij}}{\partial \ln c_{ij}} = \begin{cases} \frac{c_{ij}}{\underline{\theta}_i/(\alpha_i - 1) + c_{ij}} \in (0, 1), & \text{if } q_{ij} > q_{ji}, \\ 0, & \text{if } q_{ij} < q_{ji}, \end{cases} \quad (13)$$

and

$$\frac{\partial \ln F_{ij}}{\partial \ln c_{ij}} = \begin{cases} \frac{-2(\alpha_i - 1)c_{ij}}{(\alpha_i - 2)\underline{\theta}_i - (\alpha_i - 1)c_{ij}} < 0, & \text{if } q_{ij} > q_{ji}, \\ 0, & \text{if } q_{ij} < q_{ji}, \end{cases} \quad (14)$$

respectively.
Proof. *See Appendix A.3.*

According to equation (13), both the backhaul problem and the delivery firm's market power contribute to incomplete pass-through from the marginal shipping cost to the marginal delivery price. First, due to the backhaul problem, along the shipping direction with a lower trade flow, the marginal price p_{ij} does not depend on the marginal shipping cost, and hence the pass-through rate is $\frac{\partial \ln p_{ij}}{\partial \ln c_{ij}} = 0$. Second, along the haul with a higher trade flow, the pass-through rate is $\frac{\partial \ln p_{ij}}{\partial \ln c_{ij}} = \frac{c_{ij}}{\underline{\theta}_i/(\alpha_i-1)+c_{ij}} \in (0,1)$. In this case, as the delivery firm charges a markup, $\frac{\underline{\theta}_i}{\alpha_i-1}$, above its marginal shipping cost, only part of the change in the shipping cost is passed to the marginal price. A larger markup thus results in a more incomplete pass-through.

As shown in equation (14), the pass-through from the shipping cost to the fixed fees is non-positive. Similar to the marginal price, for the shipping direction with a lower trade flow, the fixed fee remains unchanged as the shipping cost changes, i.e., $\frac{\partial \ln F_{ij}}{\partial \ln c_{ij}} = 0$. For the shipping direction with a higher trade flow, the delivery firm *increases* the fixed fee in response to a reduced shipping cost, i.e., $\frac{\partial \ln F_{ij}}{\partial \ln c_{ij}} < 0$.

A reduction in the shipping costs, therefore, has two effects on the delivery prices: (i) it leads to a partial or zero reduction of the marginal price, and (ii) it increases or keeps the fixed fee unchanged. Both effects suggest that e-commerce sellers would experience limited welfare gains from the reduced shipping costs.

4 Model Estimation and Validation

In this section, we first outline the empirical strategy for identifying and estimating key parameters of the theoretical model using the data we have collected. We then validate the estimated model by assessing its ability to capture the essential characteristics of the data.

4.1 Estimation

To estimate our theoretical model, we need to obtain estimates of the following parameters: the Pareto distribution of willingness-to-pay θ in each city, as summarized by the shape parameter α_i and the scale parameter $\underline{\theta}_i$, as well as the marginal shipping cost c_{ij} between any pair of cities i and j. Based on the theoretical results from the previous section, we develop a block-recursive structural approach that allows us to estimate these parameters step by step.

Step 1: Inferring Shipping Flows Using the Gravity Model According to Proposition 1, the delivery prices are affected by whether the shipping direction has a greater volume than its opposite direction. We have collected data from *Cainiao* regarding the number of delivery orders between city pairs on its platform. However, since Cainiao only reports transactions in the recent few months, only 16,203 shipping routes between cities have positive transaction records. A large majority of the flows are zeros. If we use these observed trade flows, we can only estimate the model for a relatively small subset of Chinese cities. To overcome this limitation, we use the widely recognized gravity model, which also applies to e-commerce trade flows (Lendle et al., 2016; Fan et al., 2018), to infer shipping transaction volumes between the remaining cities:

$$q_{ij} = \exp(b_0 + b_1 \ln(Parcel_i) + b_2 \ln(Parcel_j) + b_3 \ln(DrivingDistance_{ij})) + \epsilon_{ij}, \tag{15}$$

where the dependent variable, q_{ij}, is the number of transactions from city i to city j observed on the Cainiao platform, $Parcel_i$ and $Parcel_j$ are the total numbers of parcel throughput of city i and j, respectively, and $DrivingDistance_{ij}$ is the driving distance between the two cities. The estimation results are shown in Table 4, with columns (1) and (2) showing results from Ordinary Least Squares (OLS) and columns (3) and (4) showing results from Poisson Pseudo Maximum Likelihood (PPML). As we can see, the number of transactions increases with the parcel throughput of the origin and destination cities but declines with the driving distance between the cities.

Table 4: Estimating Delivery Transactions Between Cities by Gravity Equation

	ln(# of transactions)		# of transactions	
	(1)	(2)	(3)	(4)
ln(parcel throughput in origin city)	0.401^a	0.402^a	0.871^a	0.820^a
	(59.02)	(59.27)	(20.43)	(32.74)
ln(parcel throughput in destination city)	0.313^a	0.297^a	0.764^a	0.700^a
	(58.58)	(53.20)	(10.66)	(14.84)
ln(driving distance)		-0.140^a		-0.634^a
		(-9.55)		(-2.94)
R-squared	0.257	0.261	–	–
No. of observations	16,203	16,203	111,220	111,220

Notes: The dependent variable is the number of transactions observed between cities on Cainiao (in logarithm in columns (1) and (2) and in levels in columns (3) and (4)). Columns (1) and (2) are estimated through OLS. Columns (3) and (4) are estimated through PPML. The numbers in the parentheses are robust standard errors. a indicates a significance level of 0.01.

Our preferred specification is column (4), where we estimate the model using PPML, which can handle a large number of zeros in our data and accounts for potential heteroscedasticity (Silva and Tenreyro, 2006). The predicted delivery service flow from city i to city j is given by

$$\hat{q}_{ij} = exp(\hat{b}_0 + \hat{b}_1 \ln(Parcel_i) + \hat{b}_2 \ln(Parcel_j) + \hat{b}_3 \ln(DrivingDistance_{ij})), \tag{16}$$

which is used in our estimation below.[15]

Step 2: Estimating the Pareto Shape Parameter For a city i, we can find the set of other cities that ship more goods to city i than from city i, i.e.,

$$\Omega_i \equiv \{k \in \mathcal{C} | \hat{q}_{ik} < \hat{q}_{ki}\}. \tag{17}$$

According to equation (9), the marginal delivery prices for shipping from city i to the cities in Ω_i do not depend on the marginal shipping costs c_{ik} due to the backhaul problem. Therefore, if we substitute the marginal price into equation (10), we find that the fixed fee satisfies the following condition:

$$F_{ik} = \frac{((\alpha_i - 2)p_{ik})^2}{2}, \; k \in \Omega_i, \tag{18}$$

which links the fixed fee to the marginal price for a given Pareto shape parameter α_i. We then invert equation (18) and obtain the following moment condition that can be used to identify α_i:[16]

$$\hat{\alpha}_i = \frac{1}{|\Omega_i|} \sum_{k \in \Omega_i} (\frac{\sqrt{2F_{ik}}}{p_{ik}} + 2), \tag{19}$$

where $|\Omega_i|$ is the cardinality of Ω_i, i.e., the number of cities that ship more parcels to city i than from city i.[17] Intuitively, given the fixed fee F_{ik}, a larger markup in the marginal price p_{ik} implies a more right-skewed demand for delivery service, i.e., a smaller Pareto shape parameter α_i. Based on the moment condition, we can then estimate the Pareto shape parameter, α_i, of the demand for delivery services in each city i, using the observed marginal prices and fixed fees across shipping routes between cities.

Using the price listings data from Cainiao, we find that the mean and the median of the estimated Pareto shape parameters across Chinese cities are 17.48 and 19.62, respectively. Appendix Figure A.3 shows the estimated distribution of the Pareto shape parameter, which features a bimodal pattern, with two peaks at around 8.5 and 22.5, respectively. Cities in western China tend to have smaller Pareto shape parameters, which center around the left peak at 8.5, while eastern and coastal Chinese cities tend to have larger Pareto shape parameters, which center around the peak on the right at 22.5.

Step 3: Estimating the Pareto Scale Parameter Similar to our strategy for identifying the Pareto shape parameter, we continue to focus on the set of cities Ω_i to identify the Pareto scale parameter

[15] In the rest of the paper, we use the hat variable "$\hat{\alpha}$" to denote the point estimate of any parameter "α."

[16] Our identification assumption is that there are unobserved i.i.d. shocks across routes for each city.

[17] If Ω_i is an empty set, α_i cannot be identified using this method. In our sample of 334 cities, Ω_i is an empty set for only one city.

for each city. From equation (9), we know that

$$\underline{\theta}_i = p_{ik}(\alpha_i - 1),\ k \in \Omega_i. \tag{20}$$

Therefore, we can obtain the following moment condition for $\underline{\theta}_i$:

$$\hat{\underline{\theta}}_i = \frac{1}{|\Omega_i|} \sum_{k \in \Omega_i} p_{ik}(\hat{\alpha}_i - 1). \tag{21}$$

Using the estimated Pareto shape parameters $\hat{\alpha}_i$ (Step 2) and the observed marginal prices p_{ik}, we can estimate the Pareto scale parameters $\underline{\theta}_i$ from the moment condition above.[18] The mean and median of the estimated Pareto scale parameters across Chinese cities are 26.72 and 30.08, respectively. Appendix Figure A.4 plots the estimated distribution of the Pareto scale parameters, which also feature a bimodal pattern, with two peaks around 17.0 and 31.5, respectively. Again, cities in western China tend to have smaller Pareto scale parameters, which center around the left peak, while eastern and coastal Chinese cities tend to have larger Pareto scale parameters, which center around the right peak.

Since both the estimated Pareto shape and scale parameters are larger in eastern and coastal regions than in western and hinterland areas, it is unclear which regions have a higher average willingness to pay for delivery services, as measured by $\bar{\theta}_i := \int_{\underline{\theta}_i}^{\infty} \theta' f_i(\theta') d\theta' = \frac{\alpha_i \underline{\theta}_i}{\alpha_i - 1}$. Even though a thinner right-tail (higher α_i) implies a lower average willingness to pay, a higher level of the demand distribution (higher $\underline{\theta}_i$), on the other hand, also leads to a higher average willingness to pay. Using the estimated Pareto shape and scale parameters, we can determine the average willingness-to-pay parameter $\bar{\theta}_i$, as shown in Appendix Figure A.5. Our results show that the eastern and coastal areas have a larger average willingness to pay for delivery services than the western and hinterland areas, notwithstanding their less right-skewed demand distributions.

Step 4: Estimating the Marginal Shipping Cost In this last step, we estimate the marginal shipping cost, c_{ij}, between any two cities i and j. We use Ω_i^c to denote the complement of Ω_i, i.e.,

$$\Omega_i^c := \{k \in \mathcal{C} | \hat{q}_{ik} > \hat{q}_{ki}\}, \tag{22}$$

which represents the set of cities that ship fewer parcels to city i than from city i. Combining equations (9) and (10), we can establish the following condition for the fixed fee, the marginal

[18] Again, the identification of $\underline{\theta}_i$ requires Ω_i to be non-empty, which is satisfied for all but one city in our sample.

price, and the marginal shipping cost:

$$F_{ij} = \frac{((\alpha_i - 2)p_{ij} - (\alpha_i - 1)c_{ij})^2}{2}, \ j \in \Omega_i^c. \tag{23}$$

We follow Donaldson (2018) to parameterize c_{ij} as a function of distance:[19]

$$\ln(c_{ij}) = \beta_0 + \beta_1 \ln(DrivingDistance_{ij}) + \epsilon_{ij}, \tag{24}$$

which assumes that the marginal shipping cost depends on the driving distance between the two cities and an error term ϵ_{ij}. We substitute equation (24) into equation (23) and estimate β_0 and β_1 through Non-linear Least Squares (NLS). The estimation result, shown in Appendix Table A.2, indicates that a 1% increase in the driving distance between two cities increases the marginal shipping cost by about 0.32%. Given the estimated parameters $\hat{\beta}_0$ and $\hat{\beta}_1$, we can obtain an estimate of the marginal shipping cost between any cities i and j:

$$\hat{c}_{ij} = exp(\hat{\beta}_0 + \hat{\beta}_1 \ln(DrivingDistance_{ij})).$$

4.2 Model Fit

We evaluate the model's performance in capturing the targeted moments, which are the trade flows and prices. We find the estimated model performs reasonably well in matching these moments.

E-commerce Delivery Flows We first examine the estimated e-commerce shipping flows between cities from Step 1 of our estimation procedure. As mentioned above, we do not have data on e-commerce trade between Chinese cities that are contemporary to our study. To validate our estimates, we compare our estimates to the estimates from Zheng et al. (2022), which provide city-level, multi-regional, input-output tables on trade flows between around 300 Chinese cities in 2017. As shown in Appendix Table A.3, our estimated e-commerce shipping flows have reasonably high correlations with their estimates. Our estimates have a correlation of 0.660 with their estimated overall trade flows (including both intermediate and final goods) and a correlation of 0.633 with their estimated trade flows of final goods.

Delivery Prices Next, we check how the model-predicted prices fit the data. Given the estimated parameters above, the predicted marginal price and fixed fee are $\hat{p}_{ij} = \frac{1}{\hat{\alpha}_i - 1}\hat{\theta}_i + \mathbb{1}_{\{\hat{q}_{ij} > \hat{q}_{ji}\}}\hat{c}_{ij}$ and

[19] Donaldson (2018) models trade costs among domestic districts in colonial India as a function of "distance along the most efficient route," which depends on the whole transportation networks. In our case, the e-commerce parcels are primarily shipped by trucks that run on the road network. Therefore, we use the shortest driving distance between cities, as obtained from the Baidu Map service, as our measure of distance.

$\hat{F}_{ij} = \frac{(\hat{\theta}_i - \hat{p}_{ij})^2}{2}$, respectively. In Table 5, we evaluate how these predicted prices compare with their counterparts in the observed data. According to rows (a) and (b), the correlation of the model-predicted marginal price with the observed marginal price in the data is 0.741, and the correlation for the fixed fee is slightly lower, at 0.700. In rows (c) and (d), we check whether the estimated model captures the salient feature of asymmetric pricing at the city-pair level. In particular, we compare the signs of the model-predicted and the observed price differences between the two shipping directions for each city pair. Our model correctly predicts the shipping direction with a higher marginal price for 71.1% of all city pairs. The success rate for predicting the higher fixed fee is slightly higher, at 74.8%. In rows (e) and (f), we examine the asymmetry in pricing at the city level by computing the difference (in logarithm) of the average shipping price for each city as an origin and as a destination. The correlation between the model predictions and the actual data of the difference in marginal price at the city level is 0.695. The correlation for the difference in the fixed fee at the city level is 0.788.

Table 5: Delivery Prices: Model versus Data

	Statistics	Definition	Value
(a)	$corr(p_{ij}, \hat{p}_{ij})$	correlation of marginal prices	0.741
(b)	$corr(F_{ij}, \hat{F}_{ij})$	correlation of fixed fees	0.700
(c)	$sign(p_{ij} - p_{ji}) = sign(\hat{p}_{ij} - \hat{p}_{ji})$	consistency in difference of marginal prices	0.711
(d)	$sign(F_{ij} - F_{ji}) = sign(\hat{F}_{ij} - \hat{F}_{ji})$	consistency in difference of fixed fees	0.748
(e)	$corr(\ln(p_{io}/p_{id}), \ln(\hat{p}_{io}/\hat{p}_{id}))$	correlation of difference in marginal prices	0.695
(f)	$corr(\ln(F_{io}/F_{id}), \ln(\hat{F}_{io}/\hat{F}_{id}))$	correlation of difference in fixed fees	0.788

Notes: This table evaluates how the estimated marginal price and fixed fee compare with the data. The variables with hats are model-predicted values. Rows (a) and (b) show the correlation between the model and the data at the city-pair level for marginal price and fixed fee, respectively. Rows (c) and (d) show the share of city pairs for which the signs of the difference between the forehaul and backhaul are the same. Row (e) shows the correlation between the model and the data for the difference (in logarithm) of the average marginal price for a city as an origin ("o") and as a destination ("d"), and row (f) shows this correlation for the fixed fee.

4.3 Validation of the Estimated Model

In this subsection, we further validate the estimated model before using it for counterfactual analyses. In particular, we examine how well it matches the non-targeted moments.

Spatial Distribution of E-Commerce Sellers and Delivery Revenues We next examine the model predictions on the measure of e-commerce sellers per city, which can be obtained from the following moment condition:

$$\hat{N}_i = \frac{1}{|\mathcal{C}| - 1} \sum_{j \neq i} \frac{\hat{q}_{ij}}{\frac{\hat{\alpha}_i}{\hat{\alpha}_i - 1}\underline{\hat{\theta}}_i - \hat{p}_{ij}}, \tag{25}$$

which is derived from equation (7). The above condition essentially states that a city has more e-commerce sellers if it has high delivery outflows, conditional on the demand from other cities and the delivery prices. Figure 4a plots the estimated spatial distribution of e-commerce sellers, which features a higher concentration of e-commerce sellers in coastal and eastern areas and areas along the Yangtze River than in other areas. Although we do not have data on the number of e-commerce sellers in each city, our estimates closely align with the findings of Huang et al. (2021), who demonstrate a similar spatial distribution of Taobao stores (see Figure 3A of their paper).

The estimated model also allows us to estimate each city's e-commerce delivery revenue. The estimated revenue for deliveries from city i to j is

$$\hat{R}_{ij} = \hat{p}_{ij}\hat{q}_{ij} + \hat{N}_i\hat{F}_{ij}\int_{\theta^*_{ij}}^{\infty}\hat{f}_i(\theta)d\theta,$$

which is assumed to accrue to the origin city i. Hence, each city i is estimated to have a total revenue of $\hat{R}_i = \sum_{j\neq i}\hat{R}_{ij}$ from inter-city parcel delivery. Figure 4b plots the estimated revenues at the city level against the delivery revenue data from the State Post Bureau (both in logarithm). The model estimates are highly consistent with the data, with a correlation of 0.947.

Table 6: Model-Predicted Price Asymmetries

	N	Mean	Std.	Median	Min	Max
Panel (a): Differences in delivery prices in and out of a city						
Marginal price	333	0.23	0.15	0.21	0.0029	0.74
Fixed fee	333	0.57	0.47	0.38	0.0083	2.13
Panel (b): Differences in delivery prices between a pair of cities						
Marginal price	55,276	0.45	0.27	0.43	0.00001	1.54
Fixed fee	55,276	0.78	0.63	0.62	0.0000003	3.39

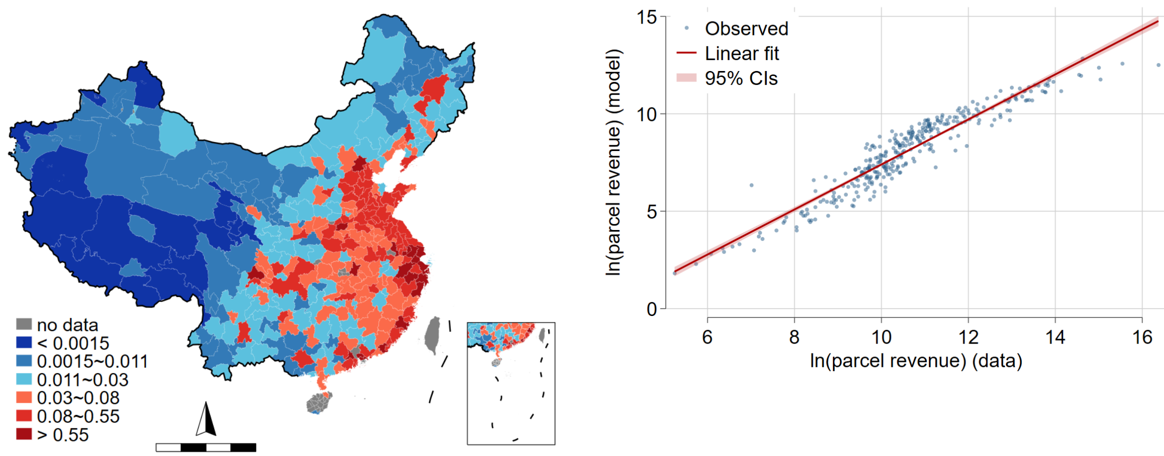
Notes: Panel (a) provides the summary statistics for the magnitude of the price differences (in logarithm) for deliveries in and out of each city, as predicted by the model. Panel (b) provides the summary statistics for the magnitude of the price differences (in logarithm) for deliveries between each pair of cities, as predicted by the model.

Stylized Facts The estimated model generates price predictions that have similar patterns as those documented in the stylized facts in Section 2.2. First, there are asymmetries in prices in and out of a city and between a pair of cities, consistent with Fact 2 in Section 2.2. In particular, Panel (a) of Table 6 presents model-predicted differences in the prices for deliveries in and out of each city (in logarithm) and presents their distributional statistics. Similar to Panel (b) of Table 2, we find substantial asymmetries in both the marginal price and the fixed fee. Moreover, consistent with the data, the model generates greater asymmetry in the fixed fee than in the marginal price. Panel (b) of Table 6 presents results at the city-pair level, and the results align with the patterns depicted in

Figure 4: Validating the Estimated Model

(a) Estimated Measure of E-commerce Sellers

(b) Revenue from Deliveries: Model v.s. Data



Notes: Figure (a) plots the estimated measure of e-commerce sellers per city. Figure (b) plots the estimated revenue from e-commerce delivery against its data counterpart and the best-fit linear line.

Table 7: Model-Predicted Two-Part Tariff Pricing

	difference in marginal price			difference in fixed fee		
	(1)	(2)	(3)	(4)	(5)	(6)
difference in parcel throughput	0.117^a	0.116^a	0.116^a	0.159^a	0.0904^a	0.0904^a
	(0.0005)	(0.0010)	(0.0010)	(0.0012)	(0.0020)	(0.0020)
difference in firm #		0.0101	0.00997		0.625^a	0.625^a
		(0.0066)	(0.0066)		(0.0147)	(0.0147)
COVID-19 policy control	No	No	Yes	No	No	Yes
R-squared	0.395	0.395	0.395	0.196	0.223	0.223
No. of observations	55276	55276	55276	55276	55276	55276

Notes: This table replicates Table 3 but replaces the dependent variables with the model-predicted prices. a indicates a significance level of 0.01.

Panel (c) of Table 2: again, there are substantial price asymmetries in the two shipping directions for each pair of cities, with greater asymmetries in the fixed fee than in the marginal price.

Finally, our model captures the relationship between delivery prices, market sizes, and market competitiveness reasonably well, as documented in Fact 3. We replicate Table 3 by conducting the regression in equation (1) using model-predicted marginal price and fixed fee as the outcome variables. The results are shown in Table 7.[20] Consistent with Fact 3, the results demonstrate that marginal prices and fixed fees out of cities with large parcel throughput tend to be higher than those out of cities with small throughput. The effect of the difference in the number of delivery

[20]The sample here is smaller than Table 3 as there is one city dropped from the estimation.

firms on the fixed fee is similar to Table 3 in columns (4) - (6), although it becomes insignificant for the marginal price in columns (1) - (3).

5 Quantitative Analysis

In this section, we first use the estimated model to quantify the importance of price discrimination versus the backhaul problem in explaining the asymmetry in delivery prices. We then conduct counterfactual analyses to examine how changes in shipping costs affect delivery prices and other outcomes.

5.1 Price Decomposition

We use equation (11) to evaluate the importance of price discrimination and the backhaul problem in driving asymmetric delivery prices. In particular, the model-predicted difference in the marginal price between the two shipping directions for a pair of cities i and j is

$$\hat{p}_{ij} - \hat{p}_{ji} = \underbrace{\frac{1}{\hat{\alpha}_i - 1}\hat{\underline{\theta}}_i - \frac{1}{\hat{\alpha}_j - 1}\hat{\underline{\theta}}_j}_{price\ \ discrimination} + \underbrace{(\mathbb{1}_{\{\hat{q}_{ij} > \hat{q}_{ji}\}} - \mathbb{1}_{\{\hat{q}_{ij} < \hat{q}_{ji}\}})\hat{c}_{ij}}_{backhaul\ \ problem}, \tag{26}$$

which decomposes the difference in marginal price into two components related to price discrimination and the backhaul problem, respectively. Similarly, from equation (12), we have

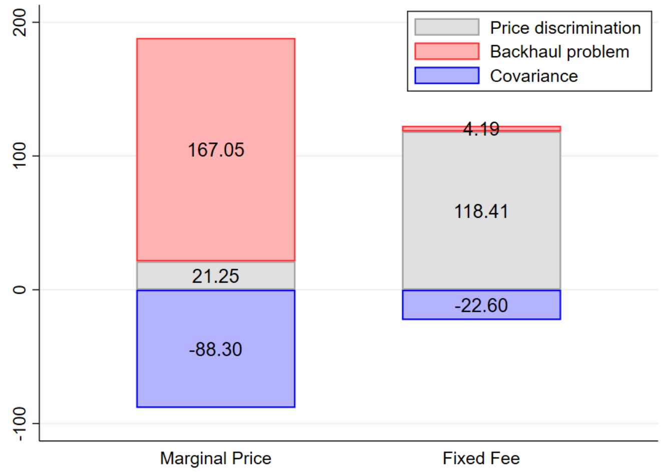
$$\frac{(\hat{F}_{ij} - \hat{F}_{ji})}{(\hat{\underline{\theta}}_i - \hat{p}_{ij} + \hat{\underline{\theta}}_j - \hat{p}_{ji})/2} = \underbrace{\frac{\hat{\alpha}_i - 2}{\hat{\alpha}_i - 1}\hat{\underline{\theta}}_i - \frac{\hat{\alpha}_j - 2}{\hat{\alpha}_j - 1}\hat{\underline{\theta}}_j}_{price\ \ discrimination} - \underbrace{(\mathbb{1}_{\{\hat{q}_{ij} > \hat{q}_{ji}\}} - \mathbb{1}_{\{\hat{q}_{ij} < \hat{q}_{ji}\}})\hat{c}_{ij}}_{backhaul\ \ problem}, \tag{27}$$

which decomposes the difference in the fixed fee between the two shipping directions.

Equations (26) and (27) enable us to evaluate the contributions of price discrimination and the backhaul problem to the observed price asymmetry through variance-covariance decomposition. Figure 5 presents the estimation results. As shown, the asymmetry in marginal prices across city pairs is primarily driven by the backhaul problem, which accounts for approximately 167% of the variation, whereas price discrimination explains only about 21%. The 88% excess variation is offset by the negative covariance between the two components. In other words, when the backhaul problem results in a higher marginal price in one shipping direction, the delivery firm tends to reduce the markup in this direction to encourage increased use of the service.

In contrast, the asymmetry in the fixed fee across city pairs is primarily driven by price discrimination, which accounts for approximately 118% of its variation, while the backhaul problem

Figure 5: Decomposition of Price Variations



Notes: This figure presents the variance-covariance decomposition for the variations of differences in both marginal price and fixed fee using equations (26) and (27), respectively. The numbers represent the percentage points of the variance of the left-hand side of (26) or (27) that can be explained by individual variances or the covariance of the components on the right-hand side of (26) or (27). The red and grey bars represent the individual variances of each component, associated with price discrimination or the backhaul problem, while the blue bar represents the covariance of the two components.

component explains only 4%. Their negative covariance between these components offsets the remaining excess variation: in shipping directions where market power leads delivery firms to charge a high fixed fee, considerations related to the backhaul problem tend to reduce it.

5.2 Counterfactual Analysis

Transportation infrastructure in many developing countries, including China, has seen significant improvements in recent decades. These improvements have brought about reductions in trade costs and increases in economic welfare. However, the welfare gains can be unevenly distributed across different geographical regions due to factors such as agglomeration and dispersion forces (Faber, 2014), imperfect competition (Atkin and Donaldson, 2015), or variations in the extent of infrastructure improvement (Allen and Arkolakis, 2022).

In this subsection, we conduct a counterfactual analysis to evaluate how the backhaul problem and price discrimination by transportation firms can lead to uneven welfare gains from reduced shipping costs. We use the estimated model to study the effect of a universal reduction in the marginal shipping cost, c_{ij}, across all city pairs. Models of perfect competition in the delivery industry without backhaul problems lead to a universal reduction in shipping prices of the same magnitude. In contrast, in our model, the presence of markups and backhaul problems leads to incomplete and asymmetric pass-through from reductions in c_{ij} to marginal prices p_{ij} and p_{ji} and fixed fees F_{ij} and F_{ji}.

Specifically, we consider a 10% reduction in the marginal shipping cost, c_{ij}, for all city pairs, say, due to a nationwide improvement in transportation infrastructure.[21] Panel (a) of Table 8 presents estimated changes in the marginal price, fixed fee, and e-commerce seller surplus at the route level as a result of the reduced shipping cost. We find that marginal price declines on average, as shown in Figure 7a, which displays the kernel density of marginal prices across shipping routes in both the baseline and counterfactual economies. It is evident that marginal prices fall following the reduction in marginal shipping costs.

Table 8: Effects of a Counterfactual Change in the Marginal Shipping Cost

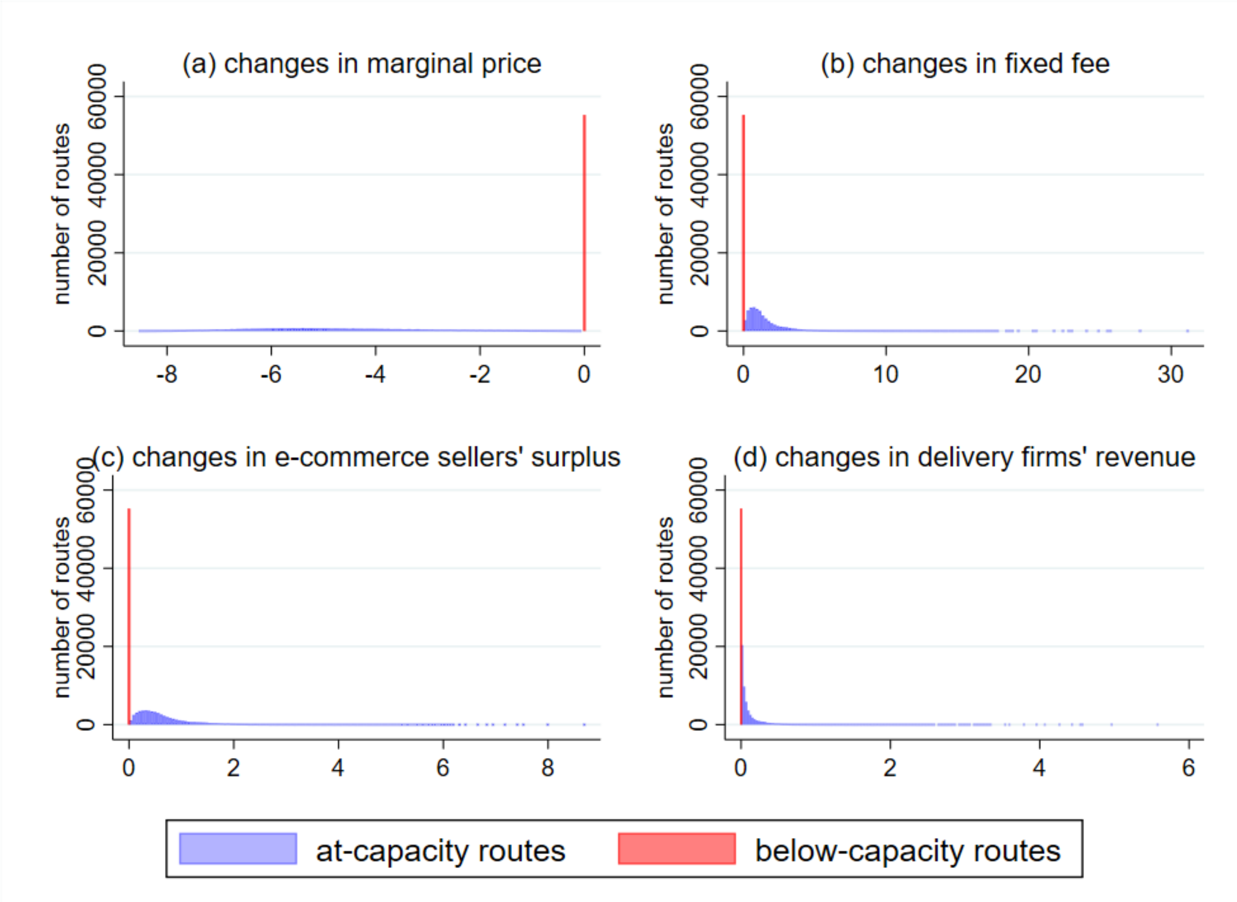
	N	Mean	Std.	Median	Min	Max
Panel (a): route-level summary statistics						
Changes in marginal prices	110552	-2.35	2.64	-0.028	-8.54	0
Changes in fixed fees	110552	0.82	1.52	0.0028	0	31.2
Changes in e-commerce sellers' surplus	110552	0.36	0.62	0.0013	0	8.72
Changes in delivery firms' revenue	110552	0.07	0.21	0	0	5.59
Panel (b): city-pair level summary statistics						
Changes in delivery firms' profit	55276	1.17	1.13	0.84	0.004	16.65
Panel (c): city-level summary statistics						
Changes in marginal prices	333	-2.35	1.40	-2.31	-5.41	0
Changes in fixed fees	333	0.82	0.46	0.75	0	2.88
Changes in e-commerce seller's surplus	333	0.34	0.17	0.33	0	0.85

Notes: This table presents the effect (in percentage points) of a universal 10% reduction in the marginal cost of deliveries. Panel (a) provides the summary statistics for changes in marginal price, fixed fee, and surplus at the route level, i.e., from one city to another. Panel (b) presents the result at the city-pair level. Panel (c) provides summary statistics at the origin city level, which are the simple averages of the route-level outcomes for each city.

However, the reduction in marginal price is smaller than the reduction in marginal cost. According to Corollary 2, there are two reasons for this incomplete pass-through. First, due to the backhaul problem, the marginal price does not change with the marginal shipping cost for routes with trade flow below capacity. This explains why half of the shipping routes exhibit no change in marginal price, as shown in the red bar in Figure 6a. Second, for routes with trade flows equal to the delivery firm's shipping capacities, the marginal price does depend on the marginal shipping cost, but the presence of markups still leads to incomplete pass-through. Among these routes, the largest decline in the marginal price is 8.54%, which is still less than the 10% reduction in

[21] See Allen and Arkolakis (2022) for a similar exercise. Note that with a reduction in c_{ij}, the values of the indicator functions $\mathbb{1}_{\{q_{ij}>q_{ji}\}}$ and $\mathbb{1}_{\{q_{ji}>q_{ij}\}}$ are not affected. To see this, suppose $q_{ij} > q_{ji}$. In this case, a reduction in c_{ij} lowers p_{ij}, increases q_{ij}, but does not affect p_{ji}. Hence, we still have $q_{ij} > q_{ji}$ at the reduced cost. In contrast, as the shipping cost c_{ij} increases, p_{ij} will increase, q_{ij} will decrease, and there can be discontinuous changes in the values of $\mathbb{1}_{\{q_{ij}>q_{ji}\}}$ and $\mathbb{1}_{\{q_{ji}>q_{ij}\}}$ as well as other outcome variables.

Figure 6: Counterfactual Effects across Routes



Notes: This figure shows the counterfactual effects of a universal 10% reduction in the marginal cost of shipping.

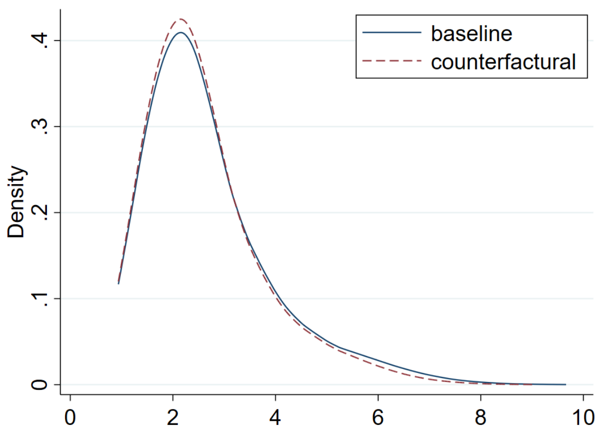
marginal cost. Overall, the average decline in marginal prices is 2.35%. Hence, the pass-through from marginal cost to marginal price is significantly less than 1 and thus incomplete.

As the marginal price declines, e-commerce sellers are incentivized to order large delivery volumes, which generates additional surpluses. Through two-part pricing, delivery firms can capture part of this surplus by raising the fixed fee, which does not affect e-commerce sellers' marginal (quantity) decisions. This is precisely what we observed in panel (a) of Table 8: the fixed fee, on average, increases by 0.82%. Changes in fixed fees also vary across routes. For routes where the marginal price remains unchanged or, equivalently, where shipping flows fall below the maximum shipping capacity for a city pair, the fixed fee does not change, as shown in Figure 6b. However, for routes where the marginal price declines, the fixed fee increases, with some increases reaching as much as 31.2%.

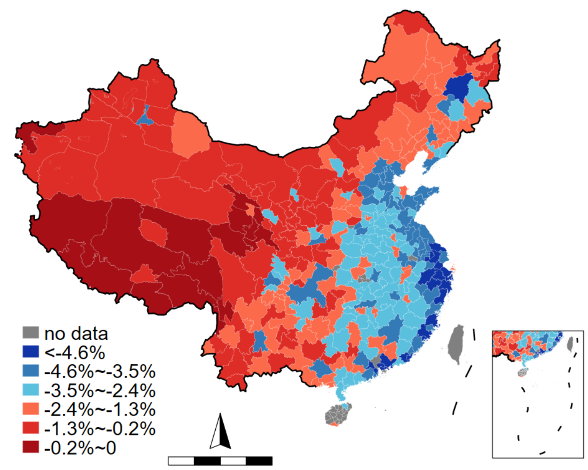
E-commerce sellers benefit from a lower average marginal price, but due to the rising fixed fee, their surpluses may not increase significantly. To quantify e-commerce sellers' welfare gains, we first aggregate the surplus across all e-commerce sellers and calculate the route-specific total

Figure 7: Marginal Price across Routes & Counterfactual Effects across Cities

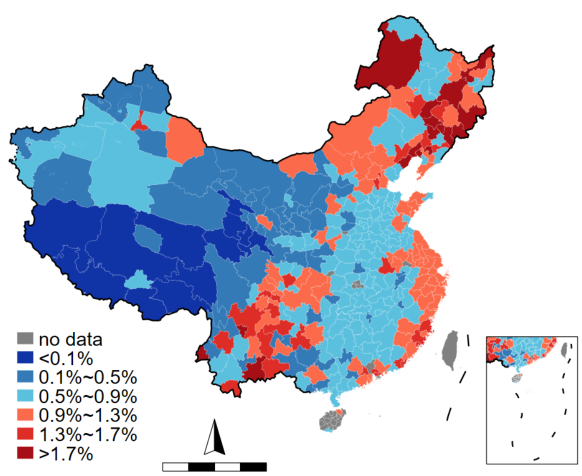
(a) Density of Marginal Price



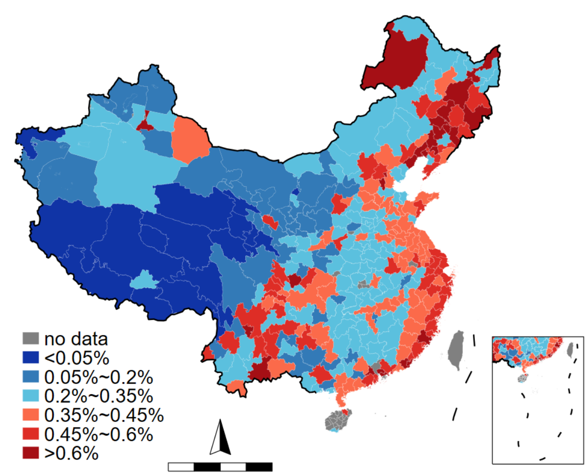
(b) Change in Marginal Price



(c) Change in Fixed Fee



(d) Change in E-commerce Seller Surplus



Notes: Figure (a) plots the kernel density plots of marginal prices in the baseline model and the counterfactual analysis. Figures (b) and (c) plot the estimated average changes in the marginal price and fixed fee for each city (as an origin for deliveries), respectively. Figure (d) plots the estimated changes in e-commerce sellers' surplus for each city.

surplus of e-commerce sellers.[22] We then calculate the estimated welfare change due to the reduction in shipping costs in our counterfactual analysis. Across all routes, the e-commerce seller surplus only increases by 0.36% on average and by a maximum of 8.72%. Due to the backhaul problem, for a given city pair, the marginal price and fixed fee remain unchanged for the route with a lower-than-capacity shipping flow, despite a decline in the marginal shipping cost, resulting in zero surplus gains (see Figure 6c).

[22]See Appendix A.4 for the derivation of the aggregate surplus at the route level and city level.

Our results also show that delivery firms' revenue and profits increase on average following the reduction of shipping costs, but the changes are uneven across routes. Depending on the demand elasticity, a falling marginal price can increase or decrease the delivery firm's revenue, while a rising fixed fee tends to increase its revenue. Panel (a) of Table 8 presents summary statistics about changes in the route-level revenue $\hat{R}_{ij}$, which increases slightly by 0.07% on average and 5.59% at maximum. But again, half of the routes experience no changes in revenue, as shown in Figure 6d. We also examine changes in delivery firms' profits, which account for both changes in revenues and reductions in shipping costs. Panel (b) of Table 8 presents the summary statistics for delivery firms' profits at the city-pair level. We find that profit increases by 1.17% on average, with a maximum increase of 16.65% and a minimum increase of just 0.004%.

The reduction in shipping costs leads to spatially uneven gains at the city level. On average, marginal prices at the origin city level decline by 2.35% (see panel (c) of Table 8). The price changes exhibit salient spatial patterns, as illustrated in Figure 7b: E-commerce sellers in eastern and coastal cities tend to experience larger declines in marginal prices compared to those in western and hinterland regions. This occurs because shipping volumes from eastern and coastal cities are typically larger than those from western and inland cities. Hence, according to equation (13), the marginal price for shipping from eastern and coastal regions tends to be responsive to changes in marginal shipping cost, whereas the marginal price for shipping to these regions is less affected by the marginal cost. As a result, the pass-through from reductions in marginal shipping costs to marginal prices is greater for eastern and coastal cities compared to western and hinterland cities.

The changes in the fixed fee are also spatially uneven, as shown in Figure 7c. Eastern and coastal cities generally experience a larger increase in fixed fees compared to western and hinterland cities. Intuitively, the fixed fee tends to increase more in cities with a significant decline in the marginal price, according to equation (10). However, since the relationship between the fixed fee and the marginal price is non-linear, some southwestern and northeastern cities also experience substantial increases in the fixed fee.

Naturally, the uneven changes in delivery prices lead to an uneven spatial distribution of e-commerce sellers' surplus. Figure 7d shows the spatial distribution of the average change in e-commerce sellers' surplus at the city level. On average, e-commerce sellers experience a 0.34% welfare gain from the reduced shipping costs (see panel (c) of Table 8). Sellers in the coastal and eastern cities tend to benefit more than those in the western and hinterland cities. However, similar to the patterns of changes in the fixed fee, some sellers in southwestern and northeastern cities also experience relatively large welfare gains.

6 Conclusion

This paper investigates the pricing strategies of e-commerce delivery firms when they face backhaul problems, offer two-part prices, and engage in third-degree price discrimination. We develop and estimate a novel theoretical model using data from the largest informatics platforms for Chinese e-commerce delivery firms. Our findings reveal that the backhaul problem and price discrimination significantly impact delivery pricing, leading to an incomplete pass-through from a reduction in the marginal shipping cost to the marginal delivery price, a negative pass-through to the fixed fee, and uneven gains across space. These findings may help policymakers develop strategies to improve the pass-through from reductions in shipping costs to delivery prices and promote equitable welfare gains across different geographical regions.

While this study offers valuable insights into the pricing of e-commerce delivery services, it is important to acknowledge its limitations. Due to data limitations, this study focuses on the interaction between e-commerce sellers and delivery firms, without considering the e-commerce consumers. Future research could integrate e-commerce consumers into the analysis to provide a more comprehensive understanding of the e-commerce delivery market.

References

Alder, S., M. Z. Song, and Z. Zhu (2023). On (Un)Congested Roads: A Quantitative Analysis of Infrastructure Investment Efficiency using Truck GPS Data. *Working Paper*.

Allen, T. (2014). Information Frictions in Trade. *Econometrica 82*(6), 2041–2083.

Allen, T. and C. Arkolakis (2022). The Welfare Effects of Transportation Infrastructure Improvements. *Review of Economic Studies 89*(6), 2911–2957.

Allen, T., D. Atkin, S. C. Cleves, and C. E. Hernandez (2024a). The Traveling Trucker Problem. In *AEA Papers and Proceedings*, Volume 114, pp. 334–339. American Economic Association.

Allen, T., D. Atkin, S. C. Cleves, and C. E. Hernandez (2024b). Trucks. *Mimeo*.

Anderson, J. E. and E. Van Wincoop (2004). Trade Costs. *Journal of Economic Literature 42*(3), 691–751.

Antràs, P., T. C. Fort, and F. Tintelnot (2017). The Margins of Global Sourcing: Theory and Evidence from US Firms. *American Economic Review 107*(9), 2514–2564.

Arkolakis, C. (2010). Market Penetration Costs and the New Consumers Margin in International Trade. *Journal of Political Economy 118*(6), 1151–1199.

Asturias, J. (2020). Endogenous Transportation Costs. *European Economic Review 123*, 103366.

Atkin, D. and D. Donaldson (2015). *Who's Getting Globalized?: The Size and Implications of Intra-national Trade Costs*. Number w21439. National Bureau of Economic Research.

Bai, J., M. Chen, J. Liu, X. Mu, and D. Y. Xu (2020). Search and information frictions on global e-commerce platforms: Evidence from aliexpress. Working Paper 28100, National Bureau of Economic Research.

Behrens, K., W. M. Brown, and T. Bougna (2018). The World is Not Yet Flat: Transport Costs Matter! *Review of Economics and Statistics 100*(4), 712–724.

Behrens, K. and P. M. Picard (2011). Transportation, Freight Rates, and Economic Geography. *Journal of International Economics 85*(2), 280–291.

Bergstrand, J. H., P. Egger, and M. Larch (2013). Gravity Redux: Estimation of Gravity-Equation Coefficients, Elasticities of Substitution, and General Equilibrium Comparative Statics under Asymmetric Bilateral Trade Costs. *Journal of International Economics 89*(1), 110–121.

Bosker, M. and E. Buringh (2020). Ice (berg) transport costs. *The Economic Journal 130*(629), 1262–1287.

Brancaccio, G., M. Kalouptsidi, and T. Papageorgiou (2020). Geography, Transportation, and Endogenous Trade Costs. *Econometrica 88*(2), 657–691.

Chao, Y. (2013). Strategic Effects of Three-Part Tariffs Under Oligopoly. *International Economic Review 54*(3), 977–1015.

Cordeau, J.-F., P. Toth, and D. Vigo (1998). A Survey of Optimization Models for Train Routing and Scheduling. *Transportation Science 32*(4), 380–404.

Coşar, A. K. and B. Demir (2016). Domestic Road Infrastructure and International Trade: Evidence from Turkey. *Journal of Development Economics 118*, 232–244.

Coşar, A. K. and B. Demir (2018). Shipping Inside the Box: Containerization and Trade. *Journal of International Economics 114*, 331–345.

Couture, V., B. Faber, Y. Gu, and L. Liu (2021). Connecting the Countryside via E-Commerce: Evidence from China. *American Economic Review: Insights 3*(1), 35–50.

Cuñat, A. and R. Zymek (2024). Bilateral Trade Imbalances. *Review of Economic Studies 91*(3), 1537–1583.

Dolfen, P., L. Einav, P. J. Klenow, B. Klopack, J. D. Levin, L. Levin, and W. Best (2023). Assessing the Gains from E-Commerce. *American Economic Journal: Macroeconomics 15*(1), 342–370.

Donaldson, D. (2018). Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. *American Economic Review 108*(4-5), 899–934.

Duranton, G., P. M. Morrow, and M. A. Turner (2014). Roads and Trade: Evidence from the US. *Review of Economic Studies 81*(2), 681–724.

Evans, W. N. and I. N. Kessides (1994). Living by the ”golden rule”: Multimarket contact in the U.S. airline industry. *The Quarterly Journal of Economics 109*(2), 341–366.

Faber, B. (2014). Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System. *Review of Economic Studies 81*(3), 1046–1070.

Fajgelbaum, P. D. and A. Khandelwal (2024, June). The value of de minimis imports. Working Paper 32607, National Bureau of Economic Research.

Fajgelbaum, P. D. and E. Schaal (2020). Optimal Transport Networks in Spatial Equilibrium. *Econometrica 88*(4), 1411–1452.

Fan, J., Y. Lu, and W. Luo (2023). Valuing Domestic Transport Infrastructure: A View from the Route Choice of Exporters. *Review of Economics and Statistics 105*(6), 1562–1579.

Fan, J., L. Tang, W. Zhu, and B. Zou (2018). The Alibaba Effect: Spatial Consumption Inequality and the Welfare Gains from E-Commerce. *Journal of International Economics 114*, 203–220.

Fitch Ratings (2020). The Blue Book of China Express Industry. Fitch Ratings Inc.

Friedt, F. L. and W. W. Wilson (2020). Trade, Transport Costs and Trade Imbalances: An Empirical Examination of International Markets and Backhauls. *Canadian Journal of Economics/Revue canadienne d'économique 53*(2), 592–636.

Gabaix, X. (2016). Power Laws in Economics: An Introduction. *Journal of Economic Perspectives 30*(1), 185–206.

Ge, C., H. Huang, Z. Wang, J. Jiang, and C. Liu (2023). Working from Home and Firm Resilience to the COVID-19 Pandemic. *Journal of Operations Management 69*(3), 450–476.

Ghani, E., A. G. Goswami, and W. R. Kerr (2016). Highway to success: The impact of the golden quadrilateral project for the location and performance of Indian manufacturing. *The Economic Journal 126*(591), 317–357.

Griva, K. and N. Vettas (2015). On Two-part Tariff Competition in a Homogeneous Product Duopoly. *International Journal of Industrial Organization 41*, 30–41.

Harris, A. and T. M. A. Nguyen (2024). Long-term relationships in the us truckload freight industry. *American Economic Journal: Microeconomics*.

Head, K., R. Jing, and J. Ries (2017). Import sourcing of chinese cities: Order versus randomness. *Journal of International Economics 105*, 119–129.

Holmberg, K., M. Joborn, and J. T. Lundgren (1998). Improved Empty Freight Car Distribution. *Transportation Science 32*(2), 163–173.

Huang, H., K. Manova, O. Perello, and F. Pisch (2024). Firm Heterogeneity and Imperfect Competition in Global Production Networks. *CEPR Discussion Paper No. 19408*.

Huang, Y., C. Lin, S. Liu, and G. Manso (2021). The Salience of Entrepreneurship: Evidence from Online Business. *Available at SSRN 3843524*.

Hummels, D., V. Lugovskyy, and A. Skiba (2009). The Trade Reducing Effects of Market Power in International Shipping. *Journal of Development Economics 89*(1), 84–97.

Ignatenko, A. (2024). Price Discrimination in International Transportation: Evidence and Implications. *Discussion Paper Series in Economics 6/2024, Norwegian School of Economics*.

Irarrazabal, A., A. Moxnes, and L. D. Opromolla (2015). The tip of the iceberg: a quantitative framework for estimating trade costs. *Review of Economics and Statistics 97*(4), 777–792.

Ishikawa, J. and N. Tarui (2018). Backfiring with Backhaul Problems: Trade and Industrial Policies with Endogenous Transport Costs. *Journal of International Economics 111*, 81–98.

Jonkeren, O., E. Demirel, J. van Ommeren, and P. Rietveld (2011). Endogenous Transport Prices and Trade Imbalance. *Journal of Economic Geography 11*(3), 509–527.

Lendle, A., M. Olarreaga, S. Schropp, and P.-L. Vézina (2016). There Goes Gravity: eBay and the Death of Distance. *The Economic Journal 126*(591), 406–441.

Levinson, M. (2016). *The Box: How the Shipping Container Made the World Smaller and the World Economy Bigger*. Princeton University Press.

Li, Y., Q. Wu, Y. Zhang, G. Huang, and H. Zhang (2023). Spatial structure of china's e-commerce express logistics network based on space of flows. *Chinese Geographical Science 33*(1), 36–50.

Limao, N. and A. J. Venables (2001). Infrastructure, Geographical Disadvantage, Transport Costs, and Trade. *The World Bank Economic Review 15*(3), 451–479.

Ma, L. and Y. Tang (2024). The Distributional Impacts of Transportation Networks in China. *Journal of International Economics 148*, 103873.

Melitz, M. J. (2003). The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity. *Econometrica 71*(6), 1695–1725.

Pitney Bowes (2023). Parcel Shipping Index 2023.

Roberts, M. J. and J. R. Tybout (1997). The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs. *American Economic Review*, 545–564.

Silva, J. S. and S. Tenreyro (2006). The Log of Gravity. *Review of Economics and Statistics 88*(4), 641–658.

Slack, B. and E. Gouvernal (2011). Container Freight Rates and the Role of Surcharges. *Journal of Transport Geography 19*(6), 1482–1489.

The World Bank and Alibaba Group (2019). *E-commerce Development: Experience from China*. Washington, D.C.: World Bank Group.

Tirole, J. (1988). *The Theory of Industrial Organization*. MIT press.

Waugh, M. E. (2010). International Trade and Income Differences. *American Economic Review 100*(5), 2093–2124.

Wilson, W. W. (1987). Transport Markets and Firm Behavior: The Backhaul Problem. In *Journal of the Transportation Research Forum*, Volume 28, pp. 325–333.

Wong, W. F. (2022). The Round Trip Effect: Endogenous Transport Costs and International Trade. *American Economic Journal: Applied Economics 14*(4), 127–166.

Yin, X. (2004). Two-part Tariff Competition in Duopoly. *International Journal of Industrial Organization 22*(6), 799–820.

Zheng, H., J. Többen, E. Dietzenbacher, D. Moran, J. Meng, D. Wang, and D. Guan (2022). Entropy-based Chinese City-level MRIO Table Framework. *Economic Systems Research 34*(4), 519–544.

Appendix

A.1 Proof of Proposition 1

For each city pair $(i,j) \in \mathcal{C}^2$, the delivery firm chooses the pricing strategy $(p_{ij}, p_{ji}, F_{ij}, F_{ji})$ to maximize its profit Π_{ij} in (8), where the aggregate quantities q_{ij} and q_{ji} are given by (7) and the threshold types $\theta^*_{ij} \geq p_{ij}$ and $\theta^*_{ji} \geq p_{ji}$ are pinned down by $S(\theta^*_{ij}, p_{ij}, F_{ij}) = 0$ and $S(\theta^*_{ji}, p_{ji}, F_{ji}) = 0$, i.e.,

$$F_{ij} = \frac{(\theta^*_{ij} - p_{ij})^2}{2} \quad \text{and} \quad F_{ji} = \frac{(\theta^*_{ji} - p_{ji})^2}{2}. \tag{A.1}$$

Since (A.1) specifies one-to-one relationships between the fixed fees and the threshold types, the delivery firm's problem is equivalent to choosing $(p_{ij}, p_{ji}, \theta^*_{ij}, \theta^*_{ji})$ to maximize the profit Π_{ij} subject to the constraints $\theta^*_{ij} \geq p_{ij}$ and $\theta^*_{ji} \geq p_{ji}$ (see the discussion in Section 3.2), where the aggregate quantities q_{ij} and q_{ji} are given by (7) and the fixed fees F_{ij} and F_{ji} are given by (A.1).

Due to the backhaul problem, the firm's delivery cost, $c_{ij}\max(q_{ij}, q_{ji})$, is not smooth in the quantities and can be equal to either $c_{ij}q_{ij}$ or $c_{ij}q_{ji}$, depending on whether $q_{ij} \geq q_{ji}$ or $q_{ij} \leq q_{ji}$. Therefore, to characterize the firms' optimal pricing strategy, we first examine two sub-problems in which the firm maximizes profit subject to the additional constraints $q_{ij} \geq q_{ji}$ and $q_{ij} \leq q_{ji}$, respectively. We then compare the optimal profits in the two sub-problems and find the delivery firm's optimal pricing strategy.

First, we restrict attention to pricing strategies that lead to $q_{ij} \geq q_{ji}$. In this case, the delivery firm's total profit is

$$\begin{aligned}
\Pi_{ij} &= \frac{1}{n_{ij}}\left[p_{ij}q_{ij} + p_{ji}q_{ji} + F_{ij}N_i\int_{\theta^*_{ij}}^{\infty} f_i(\theta')d\theta' + F_{ji}N_j\int_{\theta^*_{ji}}^{\infty} f_j(\theta')d\theta' - c_{ij}q_{ij}\right] - f_{ij} \\
&= \frac{1}{n_{ij}}\left[p_{ij}q_{ij} + p_{ji}q_{ji} - c_{ij}q_{ij} + \frac{(\theta^*_{ij} - p_{ij})^2}{2}\left(\frac{\underline{\theta}_i}{\theta^*_{ij}}\right)^{\alpha_i} N_i + \frac{(\theta^*_{ji} - p_{ji})^2}{2}\left(\frac{\underline{\theta}_j}{\theta^*_{ji}}\right)^{\alpha_j} N_j\right] - f_{ij}.
\end{aligned}$$

Differentiating the profit Π_{ij} with respect to the marginal price p_{ij}, we get

$$\begin{aligned}
\frac{\partial \Pi_{ij}}{\partial p_{ij}} &= \left[\frac{\alpha_i}{\alpha_i - 1}\frac{\underline{\theta}_i^{\alpha_i}}{\theta^{*(\alpha_i-1)}_{ij}} - p_{ij}\left(\frac{\underline{\theta}_i}{\theta^*_{ij}}\right)^{\alpha_i} - (p_{ij} - c_{ij})\left(\frac{\underline{\theta}_i}{\theta^*_{ij}}\right)^{\alpha_i} - (\theta^*_{ij} - p_{ij})\left(\frac{\underline{\theta}_i}{\theta^*_{ij}}\right)^{\alpha_i}\right]\frac{N_i}{n_{ij}} \\
&= \left(\frac{1}{\alpha_i - 1}\theta^*_{ij} + c_{ij} - p_{ij}\right)\left(\frac{\underline{\theta}_i}{\theta^*_{ij}}\right)^{\alpha_i}\frac{N_i}{n_{ij}}.
\end{aligned} \tag{A.2}$$

Similarly, the derivative of the profit Π_{ij} with respect to the marginal price p_{ji} is

$$\frac{\partial \Pi_{ij}}{\partial p_{ji}} = \left(\frac{1}{\alpha_j - 1}\theta_{ji}^* - p_{ji}\right)\left(\frac{\underline{\theta}_j}{\theta_{ji}^*}\right)^{\alpha_j}\frac{N_i}{n_{ij}}. \tag{A.3}$$

Differentiating the profit Π_{ij} with respect to the threshold type θ_{ij}^*, we get

$$\begin{aligned}\frac{\partial \Pi_{ij}}{\partial \theta_{ij}^*} &= \left[\alpha_i\left(p_{ij} - c_{ij}\right)\left(\frac{\underline{\theta}_i^{\alpha_i}}{\theta_{ij}^{*(\alpha_i+1)}}p_{ij} - \left(\frac{\underline{\theta}_i}{\theta_{ij}^*}\right)^{\alpha_i}\right) + \left(\theta_{ij}^* - p_{ij}\right)\left(\frac{\underline{\theta}_i}{\theta_{ij}^*}\right)^{\alpha_i} - \frac{\alpha_i \underline{\theta}_i^{\alpha_i}(\theta_{ij}^* - p_{ij})^2}{2\theta_{ij}^{*(\alpha_i+1)}}\right]\frac{N_i}{n_{ij}} \\ &= -\left(\frac{\underline{\theta}_i^{\alpha_i}}{\theta_{ij}^{*\alpha_i - 1}}\right)\left(p_{ij} - \theta_{ij}^*\right)\left[\frac{\alpha_i}{2}(p_{ij} - \theta_{ij}^*) - \alpha_i c_{ij}\right]\frac{N_i}{n_{ij}}. \end{aligned} \tag{A.4}$$

Similarly, the derivative of the profit Π_{ij} with respect to the threshold type θ_{ji}^* is

$$\frac{\partial \Pi_{ij}}{\partial \theta_{ji}^*} = -\left(\frac{\underline{\theta}_j^{\alpha_j}}{\theta_{ji}^{*\alpha_j - 1}}\right)\frac{\alpha_j}{2}(p_{ij} - \theta_{ji}^*)^2\frac{N_i}{n_{ij}}. \tag{A.5}$$

Consider the relaxed problem of choosing $(p_{ij}, p_{ji}, \theta_{ij}^*, \theta_{ji}^*)$ to maximize Π_{ij}, ignoring the constraints $q_{ij} \geq q_{ji}$. Since $\theta_{ij}^* \geq p_{ij}$ and $\theta_{ji}^* \geq p_{ji}$, the first-order conditions imply that the optimal solution of the relaxed problem is $\theta_{ij}^* = \underline{\theta}_{ij}$, $\theta_{ji}^* = \underline{\theta}_{ji}$,

$$p_{ij} = \frac{1}{\alpha_i - 1}\theta_{ij}^* + c_{ij}$$

and

$$p_{ji} = \frac{1}{\alpha_j - 1}\theta_{ji}^*.$$

Notice that

$$q_{ij} \geq q_{ji} \iff \frac{\theta_{ji}^{*(\alpha_j - 1)}(\theta_{ij}^* - c_{ij})}{\theta_{ij}^{*\alpha_i}} \geq \frac{\underline{\theta}_j^{\alpha_j} N_j}{\underline{\theta}_i^{\alpha_i} N_i} \iff N_i(\underline{\theta}_i - c_{ij}) \geq N_j \underline{\theta}_j,$$

$$p_{ij} \leq \theta_{ij}^* \iff c_{ij} \leq \frac{\alpha_i - 2}{\alpha_i - 1}\underline{\theta}_i,$$

and

$$p_{ji} \leq \theta_{ji}^* \iff \alpha_j \geq 2.$$

Therefore, given the parameter assumptions, when $N_i(\underline{\theta}_i - c_{ij}) \geq N_j \underline{\theta}_j$, the solution to the relaxed problem is also the solution to the firm's problem with the constraint $q_{ij} \geq q_{ji}$. Otherwise, the solution to the relaxed problem is not feasible, and hence the constraint $q_{ij} \geq q_{ji}$ must bind, i.e., $q_{ij} = q_{ji}$.

Next, we focus on pricing strategies that lead to $q_{ij} \le q_{ji}$. By symmetry, when $N_j(\underline{\theta}_j - c_{ij}) \ge N_i\underline{\theta}_i$, the firm's optimal pricing strategy (under the constraint $q_{ij} \ge q_{ji}$) is $\theta^*_{ij} = \underline{\theta}_{ij}$, $\theta^*_{ji} = \underline{\theta}_{ji}$,

$$p_{ij} = \frac{1}{\alpha_i - 1}\theta^*_{ij},$$

and

$$p_{ji} = \frac{1}{\alpha_j - 1}\theta^*_{ji} + c_{ij}.$$

Otherwise, if $N_j(\underline{\theta}_j - c_{ij}) < N_i\underline{\theta}_i$ the solution to the optimal pricing (with the additional constraint $q_{ij} \ge q_{ji}$) leads to a binding quantity constraint, i.e., $q_{ij} = q_{ji}$.

To obtain the delivery firm's (globally) optimal pricing strategy, we need to compare the profits in the two cases discussed above, i.e., pricing under the constraint $q_{ij} \ge q_{ji}$ or $q_{ji} \ge q_{ij}$, respectively. However, as shown above, under Assumption 1, the pricing strategy in one of the two cases leads to a binding quantity constraint $q_{ij} = q_{ji}$, so the pricing strategy is also feasible in the other case (when the quantity constraint is reversed). Hence, (A.1) and the analysis of the pricing strategies (of the two cases) above imply the pricing formulas in the statement of the proposition.

■

A.2 Proof of Corollary 1

Taking the difference of p_{ij} and p_{ji} in Equation (9), we have

$$p_{ij} - p_{ji} = \frac{1}{\alpha_i - 1}\underline{\theta}_i - \frac{1}{\alpha_j - 1}\underline{\theta}_j + (\mathbb{1}_{\{q_{ij} > q_{ji}\}} - \mathbb{1}_{\{q_{ij} < q_{ji}\}})c_{ij}. \tag{A.6}$$

Similarly, from Equation (10), we find that

$$\begin{aligned} F_{ij} - F_{ji} &= \frac{(\underline{\theta}_i - p_{ij})^2}{2} - \frac{(\underline{\theta}_j - p_{ji})^2}{2}, \qquad\qquad (A.7) \\ &= (\frac{\underline{\theta}_i - p_{ij} + \underline{\theta}_j - p_{ji}}{2})(\underline{\theta}_i - p_{ij} - \underline{\theta}_j + p_{ji}) \\ &= (\frac{\underline{\theta}_i - p_{ij} + \underline{\theta}_j - p_{ji}}{2})(\underline{\theta}_i - \underline{\theta}_j - \frac{1}{\alpha_i - 1}\underline{\theta}_i + \frac{1}{\alpha_j - 1}\underline{\theta}_j - (\mathbb{1}_{\{q_{ij} > q_{ji}\}} - \mathbb{1}_{\{q_{ij} < q_{ji}\}})c_{ij}) \\ &= (\frac{\underline{\theta}_i - p_{ij} + \underline{\theta}_j - p_{ji}}{2})(\frac{\alpha_i - 2}{\alpha_i - 1}\underline{\theta}_i - \frac{\alpha_j - 2}{\alpha_j - 1}\underline{\theta}_j - (\mathbb{1}_{\{q_{ij} > q_{ji}\}} - \mathbb{1}_{\{q_{ij} < q_{ji}\}})c_{ij}). \end{aligned}$$

Finally, note that, given our assumption that $0 < c_{ij} < \min\{\frac{\alpha_i - 2}{\alpha_i - 1}\underline{\theta}_i, \frac{\alpha_j - 2}{\alpha_j - 1}\underline{\theta}_j\}$, we have

$$\begin{aligned} \frac{\underline{\theta}_i - p_{ij} + \underline{\theta}_j - p_{ji}}{2} &= \frac{1}{2}(\frac{\alpha_i - 2}{\alpha_i - 1}\underline{\theta}_i - \mathbb{1}_{\{q_{ij} > q_{ji}\}}c_{ij} + \frac{\alpha_j - 2}{\alpha_j - 1}\underline{\theta}_j - \mathbb{1}_{\{q_{ji} > q_{ij}\}}c_{ji}) \\ &= \frac{1}{2}[\frac{\alpha_i - 2}{\alpha_i - 1}\underline{\theta}_i + \frac{\alpha_j - 2}{\alpha_j - 1}\underline{\theta}_j - (\mathbb{1}_{\{q_{ij} > q_{ji}\}} + \mathbb{1}_{\{q_{ji} > q_{ij}\}})c_{ij}] > 0. \end{aligned} \tag{A.8}$$

■

A.3 Proof of Corollary 2

Since the marginal price is $p_{ij} = \frac{1}{\alpha_i - 1}\underline{\theta}_i + \mathbb{1}_{\{q_{ij} > q_{ji}\}}c_{ij}$, we have

$$\begin{aligned} \frac{\partial \ln(p_{ij})}{\partial \ln(c_{ij})} &= \frac{\partial \ln(p_{ij})}{\partial p_{ij}} \frac{\partial p_{ij}}{\partial c_{ij}} \frac{\partial c_{ij}}{\partial \ln(c_{ij})} \\ &= \frac{1}{p_{ij}} \mathbb{1}_{\{q_{ij} > q_{ji}\}}c_{ij} \\ &= \frac{\mathbb{1}_{\{q_{ij} > q_{ji}\}}c_{ij}}{\frac{\underline{\theta}_i}{\alpha_i - 1} + \mathbb{1}_{\{q_{ij} > q_{ji}\}}c_{ij}}. \end{aligned} \tag{A.9}$$

The fixed fee satisfies

$$\begin{aligned} F_{ij} &= \frac{(\underline{\theta}_i - p_{ij})^2}{2} \\ &= \frac{(\underline{\theta}_i - \frac{1}{\alpha_i - 1}\underline{\theta}_i - \mathbb{1}_{\{q_{ij} > q_{ji}\}}c_{ij})^2}{2} \\ &= \frac{((\alpha_i - 2)\underline{\theta}_i - (\alpha_i - 1)\mathbb{1}_{\{q_{ij} > q_{ji}\}}c_{ij})^2}{2(\alpha_i - 1)^2}. \end{aligned} \tag{A.10}$$

Taking logarithms on both sides of the above equation:

$$\ln F_{ij} = 2\ln((\alpha_i - 2)\underline{\theta}_i - (\alpha_i - 1)\mathbb{1}_{\{q_{ij} > q_{ji}\}}c_{ij}) - \ln(2(\alpha_i - 1)^2). \tag{A.11}$$

Thus, we have

$$\begin{aligned} \frac{\partial \ln(F_{ij})}{\partial \ln(c_{ij})} &= \frac{2}{(\alpha_i - 2)\underline{\theta}_i - (\alpha_i - 1)\mathbb{1}_{\{q_{ij} > q_{ji}\}}c_{ij}} \frac{\partial((\alpha_i - 2)\underline{\theta}_i - (\alpha_i - 1)\mathbb{1}_{\{q_{ij} > q_{ji}\}}c_{ij})}{\partial \ln(c_{ij})} \\ &= \frac{-2(\alpha_i - 1)\mathbb{1}_{\{q_{ij} > q_{ji}\}}c_{ij}}{(\alpha_i - 2)\underline{\theta}_i - (\alpha_i - 1)\mathbb{1}_{\{q_{ij} > q_{ji}\}}c_{ij}}. \end{aligned} \tag{A.12}$$

Therefore, if $\mathbb{1}_{\{q_{ij}>q_{ji}\}} = 0$, we have $\frac{\partial \ln(F_{ij})}{\partial \ln(c_{ij})} = 0$. Alternatively, if $\mathbb{1}_{\{q_{ij}>q_{ji}\}} = 1$, we have $\frac{\partial \ln(F_{ij})}{\partial \ln(c_{ij})} = \frac{-2(\alpha_i-1)c_{ij}}{-(\alpha_i-1)c_{ij}+(\alpha_i-2)\underline{\theta}_i}$. Given our assumption that $0 < c_{ij} < \min\{\frac{\alpha_i-2}{\alpha_i-1}\underline{\theta}_i, \frac{\alpha_j-2}{\alpha_j-1}\underline{\theta}_j\}$, it is easy to prove that $(\alpha_i - 2)\underline{\theta}_i - (\alpha_i - 1)\mathbb{1}_{\{q_{ij}>q_{ji}\}}c_{ij} > 0$ and $\frac{\partial \ln(F_{ij})}{\partial \ln(c_{ij})} < 0$. ■

A.4 Aggregate Surplus of E-commerce Sellers

An e-commerce seller with type θ purchases $q_{ij}(\theta)$ units of delivery services to ship goods from city i to city j, which generates a surplus of

$$\begin{aligned}
S_{ij}(\theta) &= \frac{\theta^2 - (\theta - q_{ij}(\theta))^2}{2} - F_{ij} - p_{ij}q_{ij}(\theta) \\
&= \frac{\theta^2 - (\theta - (\theta - p_{ij}))^2}{2} - \frac{(\underline{\theta}_i^* - p_{ij})^2}{2} - p_{ij}(\theta - p_{ij}) \\
&= \frac{(\theta - p_{ij})^2}{2} - \frac{(\underline{\theta}_i^* - p_{ij})^2}{2}.
\end{aligned}$$

The aggregate surplus that e-commerce sellers in city i obtain from shipping goods to city j is

$$\begin{aligned}
S_{ij} &= N_i \int_{\underline{\theta}_i^*}^{\infty} \left(\frac{(\theta' - p_{ij})^2}{2} - \frac{(\underline{\theta}_i^* - p_{ij})^2}{2} \right) f_i(\theta') d\theta' \\
&= N_i \left(\frac{\alpha_i \underline{\theta}_i^{*2}}{2\alpha_i - 4} \left(\frac{\underline{\theta}_i}{\underline{\theta}_i^*} \right)^{\alpha_i} - \frac{\alpha_i p_{ij} \underline{\theta}_i^*}{\alpha_i - 1} \left(\frac{\underline{\theta}_i}{\underline{\theta}_i^*} \right)^{\alpha_i} + \frac{p_{ij}^2}{2} \left(\frac{\underline{\theta}_i}{\underline{\theta}_i^*} \right)^{\alpha_i} - \frac{(\underline{\theta}_i^* - p_{ij})^2}{2} \left(\frac{\underline{\theta}_i}{\underline{\theta}_i^*} \right)^{\alpha_i} \right).
\end{aligned}$$

Therefore, the aggregate surplus of e-commerce sellers in city i is

$$\begin{aligned}
S_i &= N_i \sum_{j \in \mathcal{C}_{-i}} \left(\int_{\underline{\theta}_i^*}^{\infty} \left(\frac{(\theta' - p_{ij})^2}{2} - \frac{(\underline{\theta}_i^* - p_{ij})^2}{2} \right) f_i(\theta') d\theta' \right) \\
&= N_i \sum_{j \in \mathcal{C}_{-i}} \left(\frac{\alpha_i \underline{\theta}_i^{*2}}{2\alpha_i - 4} \left(\frac{\underline{\theta}_i}{\underline{\theta}_i^*} \right)^{\alpha_i} - \frac{\alpha_i p_{ij} \underline{\theta}_i^*}{\alpha_i - 1} \left(\frac{\underline{\theta}_i}{\underline{\theta}_i^*} \right)^{\alpha_i} + \frac{p_{ij}^2}{2} \left(\frac{\underline{\theta}_i}{\underline{\theta}_i^*} \right)^{\alpha_i} - \frac{(\underline{\theta}_i^* - p_{ij})^2}{2} \left(\frac{\underline{\theta}_i}{\underline{\theta}_i^*} \right)^{\alpha_i} \right).
\end{aligned}$$

A.5 Oligopoly Model

In this section, we extend the model from Section 3 to a setting where e-commerce sellers endogenously select and form long-term relationships with delivery firms. We show that our baseline theoretical results and empirical findings can be interpreted as capturing equilibrium outcomes in this framework.

Specifically, suppose that a finite number $n_{ij} \in \mathbb{N}_+$ of delivery firms provide services to ship parcels for e-commerce sellers between each city pair $(i, j) \in \mathcal{C}^2$. Each firm has a constant marginal cost c_{ij} of expanding its shipping capacity between cities i and j, incurs a fixed cost of f_{ij}, and faces the backhaul problem as in the baseline model. E-commerce sellers' demand for shipping parcels between the two cities is assumed to be the same as in the baseline model.

Since e-commerce sellers typically establish long-term relationships with delivery firms and switch providers infrequently, we assume that e-commerce sellers first choose one of the n_{ij} delivery firms and then exclusively ship parcels through their chosen firm. After the e-commerce sellers each choose a delivery firm, each delivery firm sets a two-part tariff for their customer base to maximize profit. Due to the sequential nature of the model, the delivery firms do not engage in direct price competition in a Bertrand manner but instead compete indirectly through their "credible" pricing strategies in equilibrium.[23]

Proposition 2. *In any symmetric equilibrium, between cities i and j, each delivery firm faces a Pareto-shaped residual demand as in (2), and all delivery firms offer the same marginal price:*

$$p_{ij} = \begin{cases} \frac{\underline{\theta}_i}{\alpha_i - 1} + c_{ij}, & \text{if } q_{ij} > q_{ji}, \\ \frac{\underline{\theta}_i}{\alpha_i - 1}, & \text{if } q_{ij} < q_{ji}, \end{cases}$$

and the same fixed fee:

$$F_{ij} = \frac{(\underline{\theta}_i - p_{ij})^2}{2}.$$

Proof. First, in equilibrium, any individual e-commerce seller's deviation in choosing a different delivery firm does not affect the delivery firms' demand functions, as each seller is infinitesimal and therefore has a negligible impact.

In a symmetric equilibrium, the delivery firms adopt identical pricing strategies. Anticipating this, each e-commerce seller with type θ is indifferent between choosing any one of the n_{ij} delivery firms and randomly chooses a delivery firm. In a symmetric equilibrium, e-commerce sellers adopt identical randomization strategies. Thus, within each delivery firm's customer base, the distribution of the e-commerce sellers' willingness-to-pay parameter θ has the same Pareto shape as given by (2). Hence, each delivery firm's optimal pricing strategy is given by Proposition 1 and is indeed symmetric across different delivery firms. □

Proposition 2 provides a microfoundation for the assumption in our baseline model that delivery firms each face a Pareto-shaped residual demand.

[23]This approach is not only realistic in the e-commerce delivery market—where sellers typically establish long-term relationships with delivery firms—but it also helps us avoid the well-known technical challenges associated with oligopoly models involving nonlinear pricing (Chao, 2013; Griva and Vettas, 2015; Yin, 2004) and the issues of multiple equilibria in oligopoly models with the backhaul problem (Ishikawa and Tarui, 2018).

Additional Tables

Table A.1: Delivery Firms' Two-Part Tariff Pricing for year 2024

	difference in marginal price		difference in fixed fee	
	(1)	(2)	(3)	(4)
difference in parcel throughput	0.0153^a	0.0245^a	0.270^a	0.190^a
	(0.0005)	(0.0006)	(0.0017)	(0.0024)
difference in firm #		-0.104^a		0.907^a
		(0.0057)		(0.0198)
R-squared	0.0230	0.0296	0.360	0.385
No. of observations	55609	55609	55609	55609

Notes: This table estimates equation (1) using city-pair level data from 2024. The dependent variables are differences in delivery prices in the two directions of shipping between paired cities. Columns (1) - (2) examine the marginal delivery price, and columns (3)-(4) the fixed fee. Both the dependent and independent variables are in logarithm. The numbers in the parentheses are robust standard errors. a indicates a significance level of 0.01.

Table A.2: Estimation of Marginal Cost of Delivery

Dependent Variable	fixed fee
ln(driving distance)	0.316^a
	(0.0013)
constant	-1.142^a
	(0.0094)
Adjusted R-squared	0.560
No. of observations	55276

Notes: This table presents the estimation of Equation (23) with the marginal cost of delivery given by Equation (24) by the method of NLS. The variable ln(driving distance) represents the logarithm of the driving distance between two cities on the route. a indicates a significance level of 0.01.

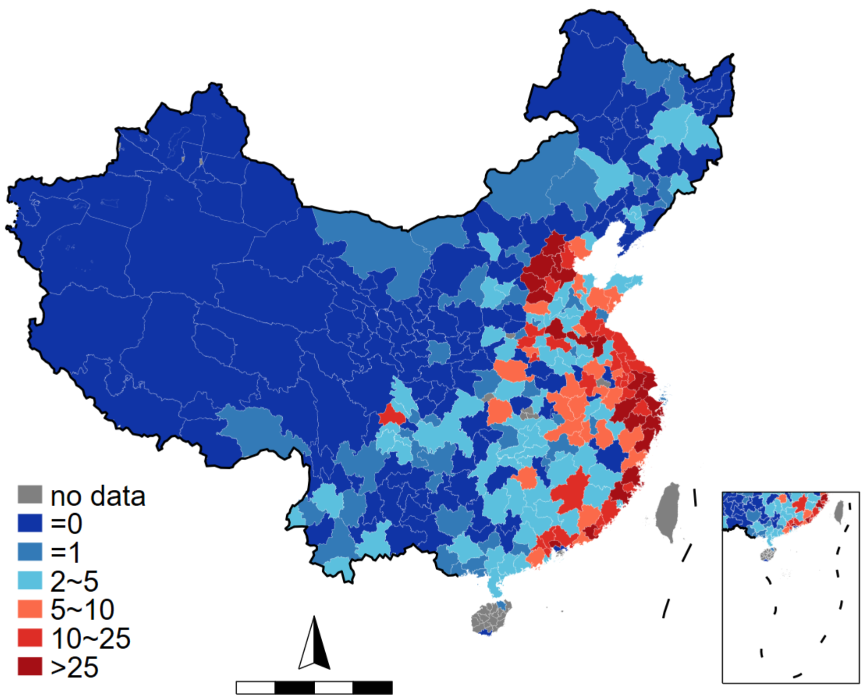
Table A.3: Correlation between Estimated Trade Flows and MRIO Data

	(1)	(2)
	$\ln(all\ trade\ flow_{ij})$	$\ln(final\ demand_{ij})$
$\ln(\hat{q}_{ij})$	0.660	0.633

Notes: This table shows the correlation between the estimated trade flows between cities and the flows extracted from city-level multi-regional input-output (MRIO) estimated by Zheng et al. (2022) for 2017. Column (1) correlates our estimated trade flows with the overall trade flows from MRIO, and column (2) with final good flows.

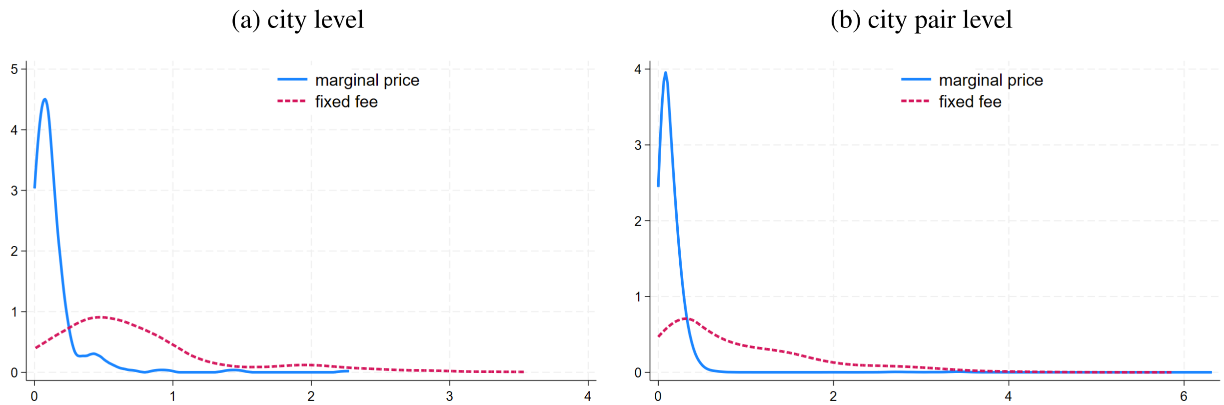
Additional Figures

Figure A.1: Number of Taobao Town per City



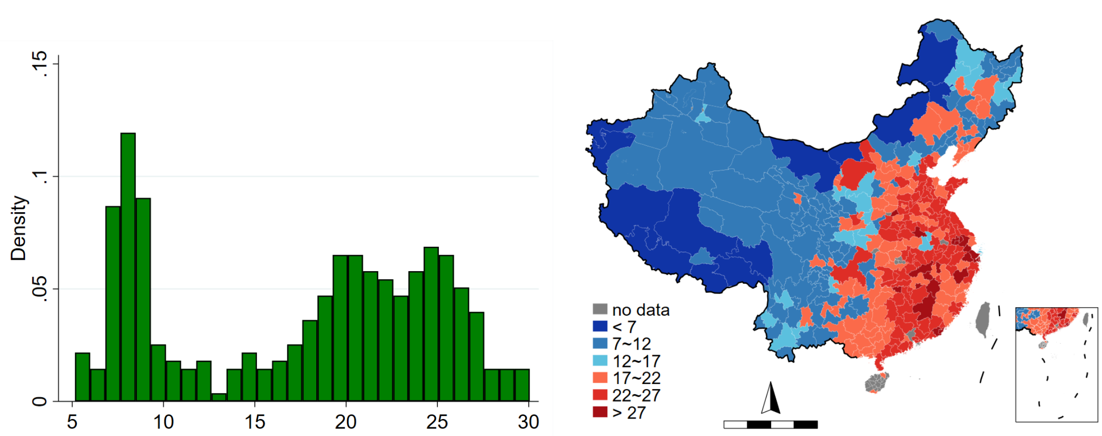
Notes: This figure plots the Number of Taobao towns per city using data from Aliresearch. AliResearch identifies a Taobao town as a town that has 3 or more active e-commerce villages, or if the annual e-commerce sales of the town reach 30 million Yuan and the number of active e-commerce stores on the Alibaba platform reaches 300 (World Bank and Alibaba Group, 2019). AliResearch has several criteria to select the villages: (1) Annual e-commerce transactions reaching 10 million Yuan; (2) the active number of e-commerce stores reaching 100 or 10% of local households. As of 2019, there were 1118 Taobao towns nationally.

Figure A.2: Kernel Density of Absolute Price Differences for year 2024



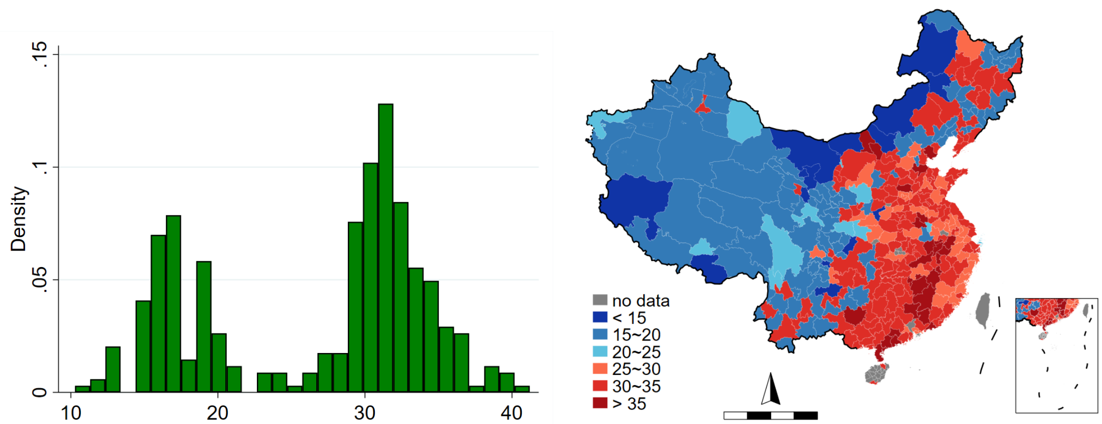
Notes: (a) plots the kernel density of absolute differences in delivery prices in and out of each city for year 2024. (b) plots the kernel density of the absolute differences in the delivery prices in the two directions of deliveries between each pair of cities in 2024.

Figure A.3: Estimated Pareto Shape



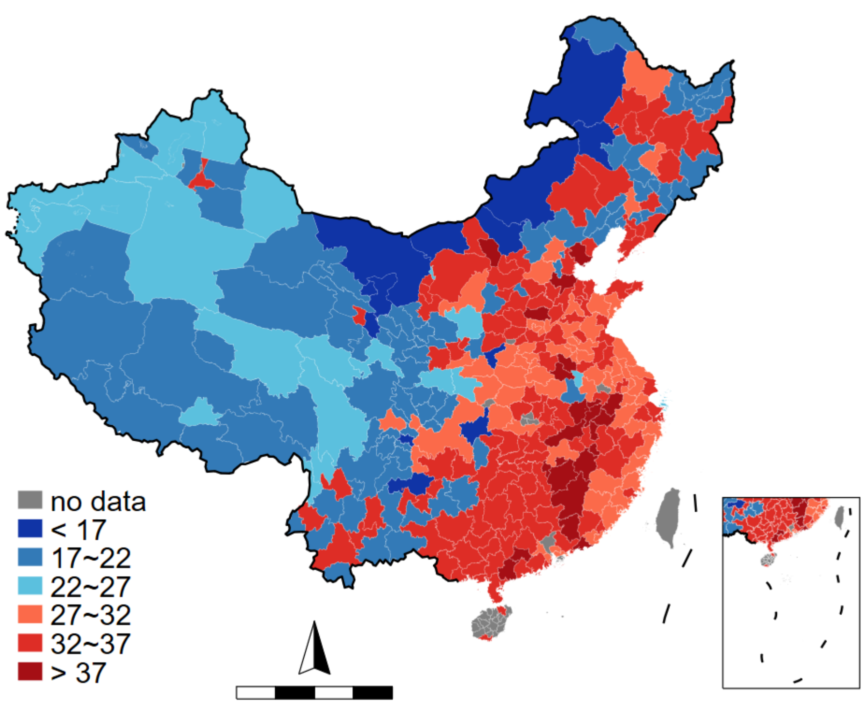
Notes: The figure on the left is a histogram of the estimated Pareto shape parameters across cities. The figure on the right is a map of each city's estimated Pareto shape parameter.

Figure A.4: Estimated Pareto Scales



Notes: The figure on the left is a histogram of the estimated Pareto scale parameters across cities. The figure on the right is a map of each city's estimated Pareto scale parameter.

Figure A.5: Average Seller Type



Notes: This figure plots the estimated average type for each city, i.e., $\frac{\hat{\alpha}_i}{\hat{\alpha}_i-1}\underline{\theta}_i$.