The China Shock and Internal Migration: Evidence from Bilateral Migration Flows*

Jaerim Choi† Yonsei University

Hyoungchul Kim‡ University of Pennsylvania

Seung Hoon Lee§ Yonsei University

March 30, 2025

Abstract

Using Korean administrative datasets spanning almost two decades and covering nearly the entire bilateral internal migration flows between local labor markets, we identify the causal impact of the China trade shock on internal migration. The trade shock affects in-migration, but not out-migration. Separating further the China trade shock into export and import shocks, we find that export expansion increases in-migration, whereas import competition reduces in-migration. By decomposing the impact of trade shock into age groups, we find that effects of trade shocks in destination are robust and statistically significant across age groups; and most pronounced for middle-aged people between the ages of 45 and 64. Finally, households with male heads are more likely to be influenced by the trade shock compared to those with female heads, due to the greater reliance of the manufacturing sector on male labor.

Keywords: China trade shock, labor adjustment, internal migration, Korea

JEL Code: F14, F16, J61, R23

*We are grateful to Robert Dekle, Andrew Greenland, David Hummels, Jay Hyun, Manho Kang, Kyong Hyun Koo, James Lake, Akira Sasahara, Andrey Stoyanov, Ariel Weinberger and other participants in Midwest International Trade Conference and Korea's Allied Economic Associations Annual Meeting for invaluable comments and suggestions. This work was supported by Yonsei University [Research Fund No. 2023-22-0442] and Yongwoon Scholarship Foundation [Grant No. 2023-11-1234]. All errors are our own.

†School of Economics, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, South Korea (e-mail: jaerimchoi@yonsei.ac.kr).

‡The Wharton School, University of Pennsylvania, 3733 Spruce Street, Philadelphia, PA 03722, USA (e-mail: hchulkim@wharton.upenn.edur).

§Corresponding author. School of Economics, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, South Korea (e-mail: seunghoonlee@yonsei.ac.kr).

1 Introduction

Recently, much attention has been paid to internal migration responses to trade liberalization, especially dealing with a puzzling phenomenon: a missing migration problem (Balsvik et al., 2015; Dix-Carneiro and Kovak, 2019; Twinam, 2022; Autor et al., 2021; Borusyak et al., 2022). Many economists have confirmed that "gains from trade" do not ensure an even distribution of the benefits arising from trade liberalization. Increased trade can allow some firms and industries a broader market for export and a lower price of intermediate goods. However, trade can also generate new challenges for others, as it can hurt certain domestic industries and their employees. Previous studies on trade liberalization have highlighted such disparate impacts of trade shocks across different regions. Specifically, they showed that regions with more intense import competition face more severe adverse effects due to trade liberalization (e.g., Autor et al., 2013).

Given the spatial disparities between winners and losers arising from trade liberalization, migration naturally emerges as a potential response to trade shocks. In order to adjust to such shocks in local labor markets, it is plausible to predict that workers will migrate from declining to more prosperous regions. However, numerous previous studies attempting to explore this pattern have found inconclusive or null results. Many related studies in Brazil, Europe, and the United States have indicated only a limited migration response to trade shocks (Autor et al., 2013; Hanlon, 2017; Balsvik et al., 2015; Donoso et al., 2015; Dix-Carneiro and Kovak, 2019), while more recent studies have noted migration responses to trade shocks (Greenland et al., 2019; Twinam, 2022). To further complicate the matter, owing to data limitations, most of the aforementioned studies employed net population changes in regions to examine labor adjustment to trade shocks. Hence, they might not have been able to comprehensively assess the impact of trade shocks on migration. While this method offers an indirect way to gauge internal migration, the measure does not fully capture gross migration flows as it fails to separate in-migration and out-migrations flow from each other. For example, the net population change will be zero if the magnitude of in-migration equals that of the out-migration. Thus, even significant migration flows might be undetected when solely examining net population changes.[1] Due to these various difficulties, the issue we refer to as the "missing migration problem" remains a puzzle in international trade.

In this paper, we aim to answer this "missing migration problem" using unique bilateral migration flows data in Korea. Starting as far back as 1996, South Korea has doc-

[1]Net population changes also include births and deaths, which further makes it hard to capture actual migration flows.

umented yearly internal migration of its whole population.[2] Notable advantages of the dataset are that (i) it records bilateral migration flows at the level of all households in South Korea, (ii) the moving of households is documented between detailed geographical units (down to regions with average population of 20,000), (iii) it includes basic demographics (age, gender) for all household members and their reason for migration. This highly detailed dataset on internal migration enables us to directly measure the impact of regional trade shocks on internal migration.

By leveraging this detailed micro-level administrative data on bilateral migration, we uncover some novel findings as to the causal impact of the China trade shock on bilateral migration flows. First and foremost, we find that the trade shock mainly affects in-migration but not out-migration. Separating the China trade shock into export and import shock shows that the export shock is associated with an increase in in-migration, while conversely, the import shock is related to a reduction in in-migration. We also demonstrate that our baseline results are robust to alternative modifications in our data or analytical methods. An interesting aspect of our results is that unlike the significant effects of trade shock in the destination, effects of trade shock in the origin appear to be insignificant. Interestingly, this null result at the origin appears to be aligned well with the past studies on the nexus between trade and internal migration where they primarily focused on push factors of trade shocks (i.e., out-migration) and found only little mobility response. In this context, our results complement the limited mobility responses to the trade shock at the origin in the literature. However, our paper goes beyond the literature by providing evidence that the trade shock affects in-migration at the destination level (i.e., pull factor).

Second, our analysis further reveals that the lack of a discernible impact on out-migration may be attributed to the heterogeneous responses observed across age groups. In preliminary analysis, we discover that the responses of younger and older heads of households to the origin trade shock are inverse: while older heads of households move away from the origin import competition and move less from the origin export opportunity, younger heads of households respond reversely. We try to answer such disparities by restricting the sample to only job-related migration and excluding major metropolitan area (capital city Seoul). This appears to nullify the previously observed heterogeneous responses to the origin trade shock across age groups, rendering all groups non-responsive to the origin trade shock. This again verifies our previous finding that trade shock appears to influence in-migration, but not out-migration.

Finally, migration responses to the trade shock are also heterogeneous across gender.

[2] As of year 2023, it has yearly migration flows data up to 2022.

The effects of both export and import shock are more pronounced for migration of male head of household. The gendered impacts of the trade shock is in line with Autor et al. (2019) where they find that Chinese import competition reduced manufacturing labor demand. As the trade shock exerts more on men's labor market outcomes, larger migration effects among male head of household are not surprising.

Our paper is related to several strands of literature on trade liberalization and its impact on local labor market and migration patterns. For example, Hanlon (2017) and Dix-Carneiro and Kovak (2019) each analyzed the effect of transitory shock in the British textile sector caused by U.S. Civil War and the effect of trade liberalization in Brazil respectively. However, these previous researches on the trade shock and its influence on local labor markets did not provide strong evidence for a substantial magnitude of mobility responses due to the onset of trade shock. Also, Facchini et al. (2019) analyzed the impact of China's integration into the world economy on internal migration in China mainland. While this study identified a notable increase in internal in-migration due to the decrease in regional trade policy uncertainty, its scope is limited due to China's stringent policies on individual internal migration decisions, known as the "Hukou system." Our research holds an advantage in studying a context where no such restrictions on individuals' internal migration decisions exist.

This study is especially related to studies on the impact of China trade shock in migration. Autor et al. (2013) discovered that the significant increase in imports from 1990 to 2007 adversely affected both employment and wages in areas susceptible to competition from China. Yet, this did not result in a notable decrease in population in these regions, implying that the geographic movement of affected workers could be limited. Building on the research of Autor et al. (2013), Greenland et al. (2019) found that the China shock indeed resulted in decreased growth rates in affected areas; however, this decline manifested with a notable lag. Other researches on China shock and migration dealt with similar results. For example, Balsvik et al. (2015) and Donoso et al. (2015) each examined the effect of China shock on Norway and Spain respectively. However, both researches found only little mobility responses. Our study significantly enriches the existing literature by leveraging unique administrative data from Korea to precisely pinpoint the previously unidentified mobility response related to the China shock. Utilizing bilateral migration data encompassing the entire population, our findings suggest that the perceived absence of a mobility response in previous literature might stem from the fact that trade shocks primarily influence in-migration, rather than the out-migration that has been the focus of past research.

Fairly recent researches on trade shock and migration we can find are Autor et al.

(2023), Twinam (2022), and Borusyak et al. (2022). First, Autor et al. (2023) found that U.S. commuting zones with higher exposure to the China trade shock experienced significant net decreases in the population of foreign-born workers. Our approach differs from theirs in that we use novel bilteral migration flows instead of using population changes to capture worker's mobility responses. We also find significant mobility patterns for native workers which was not found in Autor et al. (2023). Second, Twinam (2022) used quartz crisis in Switzerland to show that trade shock led to a rapid loss of population (i.e., out-migration). While he did find a higher level of migration responses to trade shock, our approach differs in several dimensions: (i) We use the actual bilateral migration flows rather than population changes; (ii) Our results mainly come from in-migration, whereas his results come from out-migration; (iii) While he focused on import competition, we look at both import and export shocks originating from the China trade shock; (iv) We focus on entire regions with bilateral flows data while his research focused mostly on one region. Our paper is most closely related to Borusyak et al. (2022) where they also introduce bilateral nature of location choices in migration to accurately assess the effect of trade shock on labor adjustment. Using a model of local labor markets with mobility costs and census of the Brazilian formal labor market (RAIS data), they show that past conclusions on migration can be misleading due to this bilateral nature of migration. Our paper is complementary to their paper in that our data on bilateral migration records movement for the whole population, unlike RAIS data which only covers movement for employed workers in the formal sector. By providing movement records for all population, we can also account for employees in the informal sector and unemployed individuals. Also, we add onto their research by providing a detailed information on demographics and family structure of the moving individuals. In Borusyak et al. (2022), they model worker's location choice using data on mobility cost across industries and locations. As our data contains information in the level of household and provides basic demographics (age and gender) for its members, we can introduce new channel of mobility costs induced by age and family structure.

Our paper is also related to studies examining the impact of trade shock in the context of South Korea. Kim (2006) and Choi and Kim (2015) found that trade growth with China positively influenced employment in Korea. provided a model estimating that the impact of the FTA on employment differs across industries. On the other hand, Koo and Whang (2018), concentrating solely on the manufacturing sectors, show that while the China shocks have led to job losses, they have simultaneously catalyzed job creation, thereby offsetting the impact. Our paper extends this strand of literature by examining the impact of China trade shock on the mobility responses in South Korea.

The rest of the paper is organized as follows. Section 2 explains the data and constructs the exposure to the trade shock. Section 3 discusses our empirical specification. Section 4 shows our baseline results and their robustness. Section 5 delves into the heterogeneous effects of trade shocks based on age and gender, and further explores potential reasons underlying these varied impacts. Finally, section 6 concludes.

2 Data

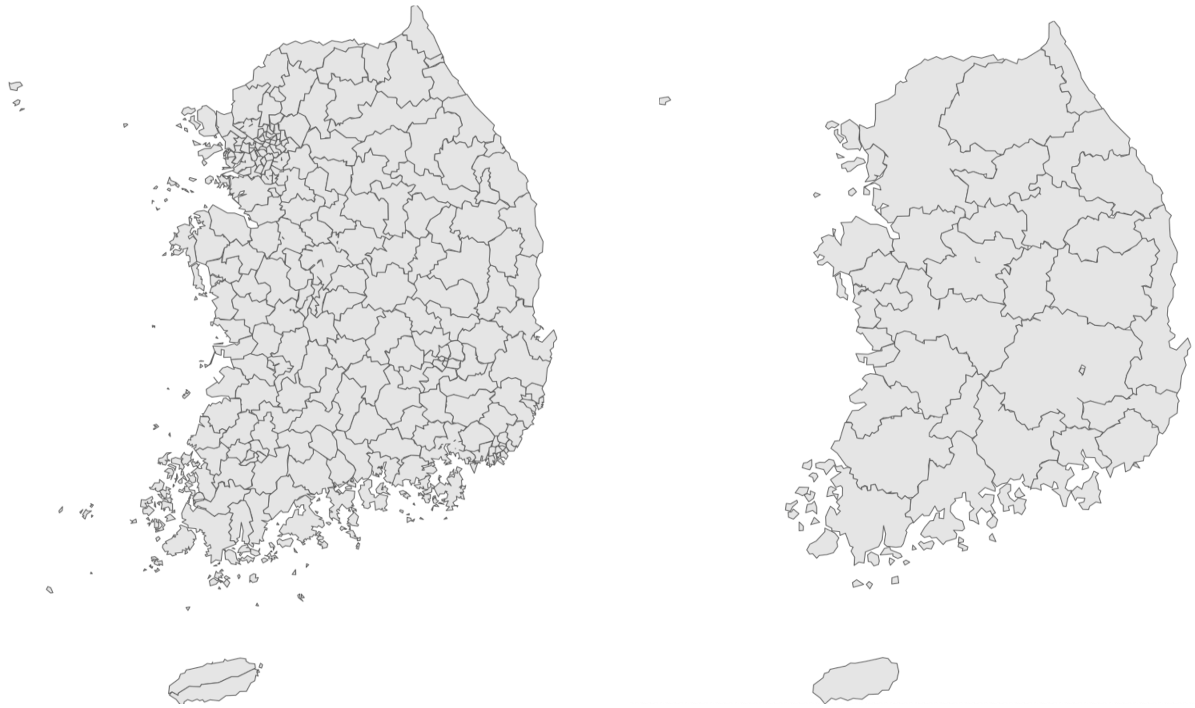
2.1 Migration Data

Our primary data set is the Internal Migration Data provided by Statistics Korea, which comprises administrative records of all bilateral migrations within South Korea. As mandated by law, Korean residents are required to report their new addresses when they move their residential locations. This dataset is provided at the household level and contains information about the age and the gender of each household member, as well as the primary reason for migration.

We aggregate this household-level data to construct bilateral migration flows across the "si-gun-gu" administrative units — which will be referred to as "districts" throughout the paper. While the dataset offers location information at an even finer unit called "dong", we utilize "si-gun-gu" as our primary unit of analysis because it is analogous to a county in the U.S. context. Although its average population size is approximately 20,000, roughly twice as large as that of a county, both geographic units represent the smallest administrative units responsible for essential government services like education, fire departments, and police services, etc.

We utilize these 226 districts for our migration analysis, while the exposure to the trade shock is calculated at the level of local labor market.[3] To be consistent with commuting zone in the U.S., we employ the concept of Travel-to-Working Areas provided by Lee and Lee (2015), who establish commuting zones in Korea based on commuting flows. Following their methodology, the South Korean districts are grouped into 33 regions. Figures 1(a) and 1(b) show the map of South Korea, delineating its districts and commuting zones, respectively.

[3]The total number of districts can differ annually due to changes in administrative divisions. We minimize this difference by keeping most of the cities that were both present in between 2001 and 2020. The detailed process is explained in the Appendix C.



(a) Districts (b) Commuting Zones

Figure 1: Map of South Korea

Notes: Each figure shows the districts (si-gun-gu) and commuting zones of South Korea respectively. There are total 226 districts and 33 commuting zones.

2.2 Trade Data

We use UN Comtrade Database on South Korea imports and exports at the 6-digit Harmonized System (HS) product level to construct the trade shock using imports from China to South Korea and exports from South Korea to China. We use annual imports from and exports to China for the years 2001, 2010, and 2019 to construct the change in the level of trade volume. Table 1 shows the trade value of Korea imports from China and exports to China for year 2001, 2010, and 2019. This table demonstrates a significant surge in trade volume between South Korea and China from 2001 to 2010, followed by a period of relative stability after 2010.

2.3 Measuring exposure to the Trade Shock

To examine the effect of trade shock at the regional level, we need to distribute Korea's trade volumes with China to each commuting zones in South Korea. We mainly construct this local labor market exposure à la Autor et al. (2013). However, there are some differences in our specification that enriches past studies on the China trade shock. First,

Table 1: Value of Trades for Korea (2001, 2010, 2019)

	Trade with China (in billions USD)		
Year	Export	Import	Balance
2001	25.6	18.7	6.9
2010	136.5	83.6	52.9
2019	136.2	107.2	29.0

Notes: All values of export and import are from Comtrade dataset. All values are adjusted to year 2019 in billions USD.

in addition to import competition channel, we also consider an export expansion margin (e.g., Dauth et al., 2014; Choi and Xu, 2020). The China trade shock's direct influence can manifest in two opposing forces within South Korean manufacturing, particularly through the interconnection of the East Asian global value chain. These forces are the export creating channel (which increases demands for Korean products by Chinese firms) and the conventional import competition channel. In this case, failing to consider both channels may result in a biased understanding of the effects of the China trade shock.[4]

In addition to considering the export creating channel, we decompose the trade shock by origin and destination. Past studies on the China trade shock and its impact on migration focused primarily on the origin trade shock using net population changes in each region (Greenland et al., 2019; Twinam, 2022). Our specification extends these studies by precisely separating this trade shock into origin and destination shocks. Moreover, we use internal migration flows between regions, instead of the net population change in a region. This allows us to more accurately assess the impact of both origin and destination trade shocks on internal migration.

In summary, we distribute Chinese imports and exports to each commuting zone based on their respective shares to the national industry employment. For convenience, we present below the equation only for import competition as export opportunity is constructed analogously:

$$\Delta IP_{ct} = \sum_{i} \frac{L_{ict}}{L_{it}} \cdot \frac{\Delta M_{it}}{L_{ct}} \tag{1}$$

where c, i, and t each refers to commuting zone, industry, and time, respectively. L_{ict} is

[4]The importance of accounting for this additional dimension of exposure in countries like Korea (and problem of not accounting for it) is discussed in more detail in the cited examples.

the number of employment at the start of period (i.e., year 2001) in industry i in region c and L_{it} refers to the total number of employment at the start of period (i.e., year 2001) in industry i in South Korea. L_{ct} is the total number of employment in region c at the start of period (i.e., year 2001). ΔM_{it} is the observed change in total Korean imports from China in industry i between the start of the period and the end of the period.

The source for the trade data ΔM_{it} is described in Section 2.2. For employment data, we use Korea's Census on Establishments. It is an annual survey consisting of about 4.4 million establishments that have one or more employees and are doing business in Korea. This data provides information on employee counts, industry codes, and locations for all establishments within South Korea.

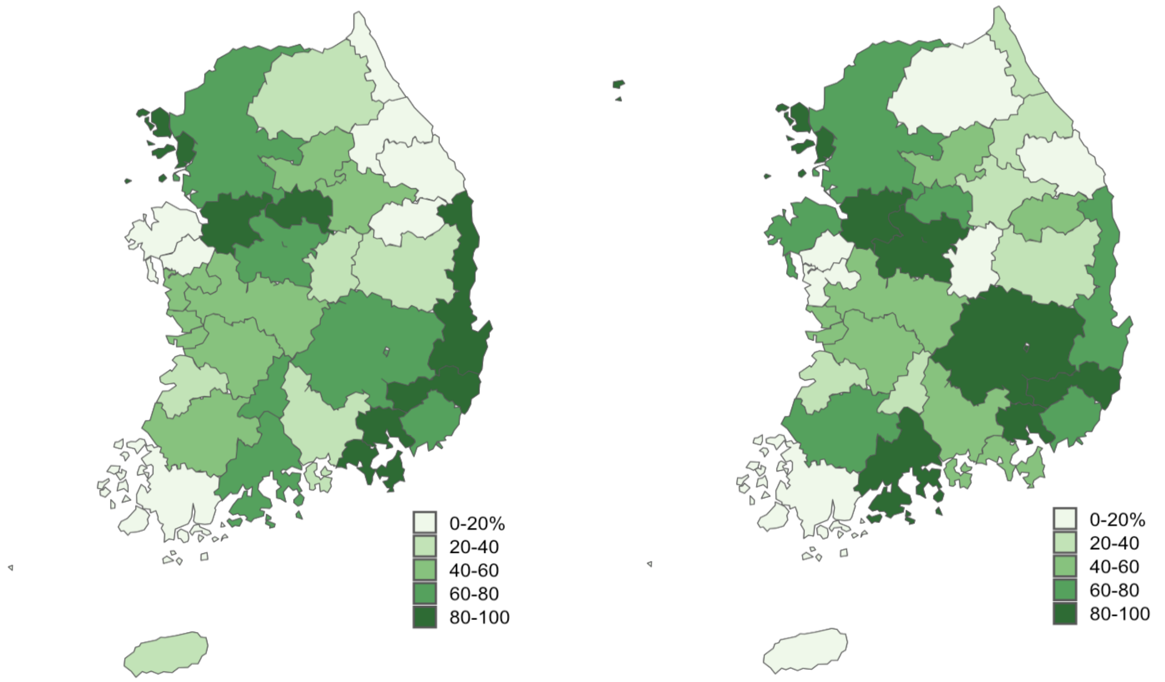
Our dataset covers the years from 2001 to 2019. We segment these 19 years into two distinct periods denoted as $t = 1$ and $t = 2$. Here, $t = 1$ corresponds to the timeframe from 2001 to 2010, while $t = 2$ corresponds to the period spanning 2010 to 2019. With the data described above, we can compute each term in equation (1). For period $t = 1$ (and similarly for $t = 2$), ΔM_{it} refers to the change in total Korean imports from China in industry i" between the years 2001 and 2010 (or between 2010 and 2019). In all of our analysis, units of imports and export exposure are in $1,000 USD. Also, value of imports and exports in year 2001 and 2010 are all adjusted to 2019 to account for inflation. For export exposure, all variables remain same, with the exception of ΔM_{it}, which is now adjusted to represent Korean exports to China. We denote export exposure in commuting zone c by ΔEX_{ct}. In Figure 2, we visualize the distribution of trade shocks in the local labor market for period 1 (2001-2010) by mapping import and export exposures across commuting zones in quantile.

Similar to past literature on estimating causal effect of China trade shock, our estimation may suffer from endogeneity issues. In order to alleviate this concern, we construct following instrumental variable à la Autor et al. (2013) using Japanese imports from China and Japanese exports to China[5]:

$$\Delta IP_{jct} = \sum_i \frac{L_{ict-1}}{L_{it-1}} \cdot \frac{\Delta M_{jit}}{L_{ct}}. \tag{2}$$

The instrumental variable differs from the model in equation (1) in two key aspects. First, we use Japanese imports from China (ΔM_{jit}) instead of Korea. Second, we use lagged employment levels $\left(\frac{L_{ict-1}}{L_{it-1}}\right)$. To be specific, we use the year 1999 for our lagged em-

[5]This setup aligns with the identification strategy used in Choi and Xu (2020), which also analyze the impact of the China Shock in the case of South Korea.



(a) Import shock

(b) Export shock

Figure 2: Map of Trade shocks (2001-2010)

Notes: Each figure plots the import and export shock for local labor market in South Korea. For visualization, values are give in quantile.

ployment levels as concurrent employment levels may be affected by anticipated China trade. The instrumental variable for export exposure is also analogous to that of import exposure and denoted by ΔEX_{jit}.

One of the key challenges in constructing local exposure to trade shock is constructing ΔM_{it} and ΔM_{jit}: we need to concord the trade imports and exports in 6-digit HS product level to 5-digit Korean Standard Industrial Classification (KSIC) industry level in South Korea. By constructing and following consistent methodology, we successfully concord product level trade shock into industry level. In the Appendix B, we provide a more thorough process of this crosswalk from HS product codes to KSIC industry codes.

3 Empirical Specification

After constructing local exposure to trade shock and its instrument variable, we specify the following 2SLS regressions with multiple endogenous variables:

$$\log(\text{Migrate}_{odt}) = \beta_0 + \beta_1 (\Delta EX_{et} - \Delta IP_{et}) + \beta_2 (\Delta EX_{st} - \Delta IP_{st}) + \beta_3 \log(\tau_{od}) + \beta_4 \log(\text{pop}_{dt}) + \beta_5 \log(\text{pop}_{ot}) + X'_{ot}\Gamma_1 + X'_{dt}\Gamma_2 + \varepsilon_{odt}, \quad (3)$$

$$\log(\text{Migrate}_{odt}) = \gamma_0 + \gamma_1 \Delta EX_{et} + \gamma_2 \Delta IP_{et} + \gamma_3 \Delta EX_{st} + \gamma_4 \Delta IP_{st} + \gamma_5 \log(\tau_{od}) + \gamma_6 \log(\text{pop}_{dt}) + \gamma_7 \log(\text{pop}_{ot}) + X'_{ot}\Gamma_1 + X'_{dt}\Gamma_2 + \varepsilon_{odt} \quad (4)$$

where $\log(\text{Migrate}_{odt})$ refers to the logarithmic value of the cumulative number of migrants from district o to district d in period t. Note that for our baseline regressions, we only consider bilateral migration where migrants move across two different commuting zones. This is mainly because our local exposure to trade shock is constructed in the level of commuting zone. Two terms $(\Delta EX_{et} - \Delta IP_{et})$ and $(\Delta EX_{st} - \Delta IP_{st})$ in equation (3) represent the relative exposure of Chinese export opportunities to import penetration in the origin commuting zone s and destination commuting zone e, respectively. A higher value indicates that, on balance, the origin o experienced more benefits from export opportunities compared to import penetration. On the other hand, we include all four endogenous variables separately in equation (4). τ_{od} refers to the distance (in m) between the origin and the destination district. pop_{ot} and pop_{dt} indicate origin and destination district's population at the start of the period respectively. X_{ot} and X_{dt} are set of region-specific control variables such as origin and destination share of manufacturing employment, origin and destination share of female employment, and origin and destination share of population with college education.[6]

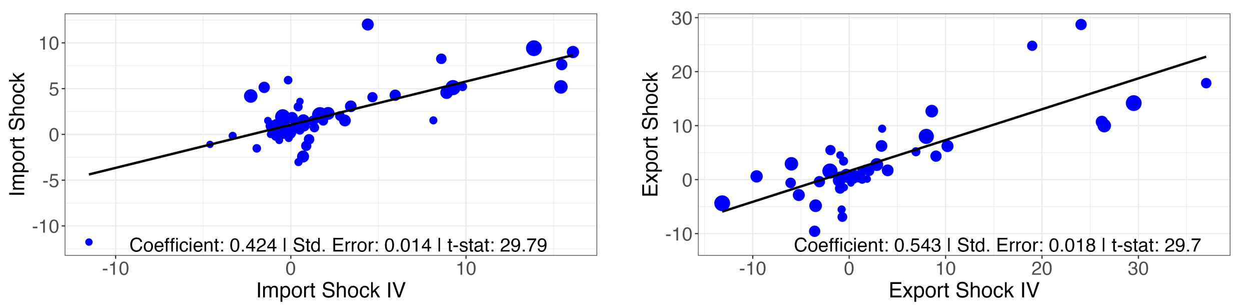
Our coefficients of interest are β_1, β_2, γ_1 through γ_4. β_1 and β_2 each captures the impact of relative exposure of Chinese export opportunities to import penetration in region d on relative bilateral migration flows between regions o and d, and the impact of relative exposure of Chinese export opportunities to import penetration in region o on bilateral migration flows between regions o and d respectively. On the other hand, γ_1 through γ_4 each captures the impact of exposure of Chinese export (import) shock in region s (e) on bilateral migration flows between region o and d. Note that in the baseline regression, all of our models are weighted by the origin population at the start of the period (2001) and standard error is clustered at the origin district (region o) level.

[6] See Section A.1 for a description of the data pertaining to commuting zone characteristics.

4 Results

4.1 First-stage Results

Figure 3 plots the correlation between our endogenous variables (the actual changes in import and export exposure in 2001-2010 and 2010-2019) on the vertical axis and their instruments, both in 1,000 USD. Each circle in the plots denotes a district. The size of the observation indicates the weight based on the initial period population in 2001. As our trade shock is defined in the level of commuting zone, we have several overlapping values. Also, we only present the plots for import and export dimension because trade shock for the same region does not differ by destination or origin.



(a) Import (b) Export

Figure 3: First-stage scatter plot

Notes: Both figures show the scatter plot for actual local exposure to trade shock and their instruments. The size of the observations indicate the weight of each district which is the population in the initial period 2001.

The scatter plots for both import and export shocks show that there is significant variation for the instrument across commuting zones. We can also verify that there are strong correlation between the endogenous variables and their instruments. We also provide basic weighted regression results for each 1st stage with no controls. The estimate yields a coefficient value of 0.424 for imports, with a t-value of 29.79, and for exports, the coefficient is 0.543 with a t-value of 29.7. This suggests that a 10 percentage point rise in the import shock instrument corresponds to a 4.24 percentage point growth in import shock. Similarly, a 10 percentage point growth in the export shock instrument relates to a 5.43 percentage point increase in export shock.

We also show our full first-stage regression results in the Appendix. Table A.2 and A.3 show our results for equations (3) and (4). Both tables show that correlations for the instrument variables and endogenous variables are highly statistically significant. They also show that instrument variable has sufficient strength, as is evident from the large first stage F-statistics ranging from 32.7 to 114,341.6.

4.2 Baseline Results

We now move onto our baseline regression results. Table 2 shows the 2SLS estimates of the impact of China trade shock on internal migration. The first column shows the regression result for equation (3) and the rest two columns are regression specifications following equation (4). Note that we added period fixed effects (two period) for all of our regression specifications. Also, for column 3, we added destination and origin region (commuting zone) fixed effects.[7]

For all three columns, higher import competition in the commuting zone of the destination seems to decrease in-migration. On the other hand, higher export opportunity in the commuting zone of the destination seems to increase in-migration. The first column shows that relative exposure of Chinese export opportunities to import penetration in the destination positively affects the in-migration from district o to district d. The estimate for the relative exposure in the destination side is 0.009, which indicates that $1,000 rise in a destination commuting zone's relative exposure per worker increases the number of migrants by approximately 0.9%. The second and third columns also reveal that the exposure of import competition (export opportunity) significantly hindered (encouraged) in-migration. For example, the estimate for the destination import shock and export shock were -0.119 and 0.04 respectively. This indicates that $1,000 rise in a destination commuting zone's exposure of import competition (export opportunity) per worker decreases (increases) the number of migrants by approximately 11.9% (4.0%).

An interesting aspect of our results is that China trade shock only seems to affect in-migration. Unlike the effects of destination trade shocks, the effects induced by origin trade shocks seem to be relatively small and insignificant. Such asymmetry in trade shocks between origin and destination could be a reason why previous studies have found only minor or insignificant effects of trade liberalization on migration. Most studies on internal migration on trade liberalization (especially China shock) focused on the impact of region's import competition on net population changes (Autor et al., 2013; Greenland et al., 2019; Balsvik et al., 2015; Donoso et al., 2015). As import competition is a push factor that drives people from the current region, this can be understood as estimating

[7] Note that there are some zero values in dependent variables that are dropped in our regression. However, as we are using cumulative migration data of 10 years for each period, such cases are very rare. In Table 2, there are only 70 cases out of 46,664 observations in each period (In total only 140 observations dropped out of 93,328 observations).

Table 2: Baseline 2SLS Results

	Dependent variable: log(migrate)		
	(1)	(2)	(3)
$(\Delta EX_{et} - \Delta IP_{et})$	0.009^{***}		
	(0.001)		
$(\Delta EX_{st} - \Delta IP_{st})$	0.003		
	(0.005)		
ΔIP_{et}		-0.044^{***}	-0.119^{***}
		(0.008)	(0.018)
ΔEX_{et}		0.015^{***}	0.040^{***}
		(0.002)	(0.006)
ΔIP_{st}		-0.056	-0.016
		(0.039)	(0.012)
ΔEX_{st}		0.019	0.005
		(0.014)	(0.005)
Controls	yes	yes	yes
Period FEs	yes	yes	yes
Region FEs	no	no	yes
Num. of Period	2	2	2
Num. Obs.	93188	93188	93188
R^2	0.772	0.767	0.804

Notes. All regressions in the table include same set of control variables mentioned in our empirical specification section. For the region fixed effects, we include origin commuting zone fixed effects and destination commuting zone fixed effects. Regression estimates are weighted by the initial population in the year 2001. Standard errors are clustered at the district level; * p < 0.10, ** p < 0.05, *** p < 0.01.

the effect of origin import penetration on out-migration in our specification. Considering this fact, it might be that previous studies could not pin down the effect of trade shock on migration because trade shock mainly affects in-migration. In this case, our results can shed new light on the missing migration problem by discovering the hidden effect of trade shocks on migration. Also, it is worth emphasizing that we measure migration responses more accurately than the previous studies that typically leverage net population changes. In Section 5, we will further analyze possible sources for such asymmetric results between destination and origin trade shocks.

4.3 Robustness Checks

Now we briefly examine if our baselines results are robust to particular changes in the data or specifications. Note that for brevity, we put all regression tables for our robustness checks in the Appendix.

We first investigate migration patterns of households led by working-age individuals (ages 25 to 64). Table A.4 demonstrates that the migration impacts of trade shocks are significant, with the magnitude of these impacts even larger for our restricted age sample, ensuring the results are not biased by samples outside working age. Interesting thing to note is that there is again a clear asymmetry: while trade shocks at the destination strongly influence migration, the origin trade shock's effects are more ambiguous or negligible.

Our initial findings were based on a two-period specification: 2001-2010 and 2010-2019. However, to ensure that the construction of these periods, especially the inclusion of 2020 (the onset of COVID-19), does not lead to significant bias, we also analyzed a single period (2001-2019). Table A.5 confirms consistency in our findings, regardless of period specification.

Given that people migrate for various reasons, we revisit our findings by examining only job-related migration. Table A.6 validates our earlier findings, emphasizing that employment-related migrations are indeed influenced by trade shocks.

In South Korea, capital city Seoul constitutes the primary metropolitan area, accounting for nearly 20% of the nation's population. South Korea also has a province named Jeju-do which is an island area separated from the mainland. As these two areas possess unique regional characteristics in Korea, there is a chance that our baseline results are mainly driven by them. To address this concern, we re-estimate our baseline results by excluding Seoul and Jeju-do. Table A.7 shows the results. The results show that the estimates are still consistent and align with the results in our baseline analysis. This again shows that our baselines results are not mainly driven by certain regions in South Korea.

We also conducted a falsification test to address potential confounding migration trends before the "China shock." This is done by regressing past migration on future trade exposures. Table A.8 showed minimal evidence for such trends, suggesting our results robustly reflect the actual impacts of trade shocks on migration.

5 Heterogeneous Effects of the Trade Shock

We now further investigate heterogeneous impacts of the trade shock on internal migration flows by exploring across age and gender groups. By dividing migration into various dimensions, we can gain a more comprehensive understanding of which specific groups are most affected and under what conditions. This method allows us to pinpoint the link between the trade shock and migration flows, thus providing a deeper understanding of mechanisms. Additionally, differentiating between age and gender can shed light on potential social or economic implications that may not be immediately evident when looking at the overall migration trends.

5.1 Age Effect

One of the puzzling phenomena in our baseline results is that unlike the destination trade shock, the effect of origin trade shock on migration is unclear and insignificant. We strive to investigate this null result by breaking down the effects of the trade shock on migration across various age groups.

To accomplish this, we specify the following regression model:

$$\log(\mathrm{Migrate}_{\mathrm{odjt}}) = \alpha_0 + \beta_j \sum_j \Delta EX_{et} \times 1\{age = j\} + \gamma_j \sum_j \Delta IP_{et} \times 1\{age = j\} + \delta_j \sum_j \Delta EX_{st} \times 1\{age = j\} + \theta_j \sum_j \Delta IP_{st} \times 1\{age = j\} + X'_{ojt}\Gamma_1 + X'_{djt}\Gamma_2 + \varepsilon_{odjt} \quad (5)$$

where j denotes an age group and $j \in \{\text{“}25-34\text{”}, \text{“}35-44\text{”}, \text{“}45-54\text{”}, \text{“}55-64\text{”}, \text{“}65+\text{”}\}$. We set migration of age 65 or more (“65+”) as a baseline group. To be specific, equation (5) is similar to equation (4) except that we now stack five different age group observations in a single regression framework.[8] We have included period by age group fixed effects.

Table A.9 in the Appendix shows the results of the full estimates. The difference between columns (1) and (2) is that we have included region fixed effects in column (2). Since there are too many variables in Table A.9, we plot the value of estimates by two dimensions: Destination - Origin, and Import - Export. This is to clearly examine the effects of trade shocks by different age groups. For all age groups, we added coefficient values for the age groups in Table A.9 to the base estimates of age 65 or more to compare the magnitude of the trade shock effects.

Figure 4 illustrates the results for the destination trade shock. At first glance, it is important to highlight that the effect of trade shock at the destination on migration remains notably consistent across different age groups. For all age groups, the destination export shock increases in-migration, whereas the destination import shock decreases in-migration. However, this result by age groups unveils a new, intriguing pattern. Both impacts for destination trade shock are most pronounced for middle-aged individuals,

[8] To be more specific, we decompose the total migration into 5 age groups as follows: (i) Age 25-34; (ii) Age 35-44; (iii) Age 45-54; (iv) Age 55-64; and (v) Age 65 or more.

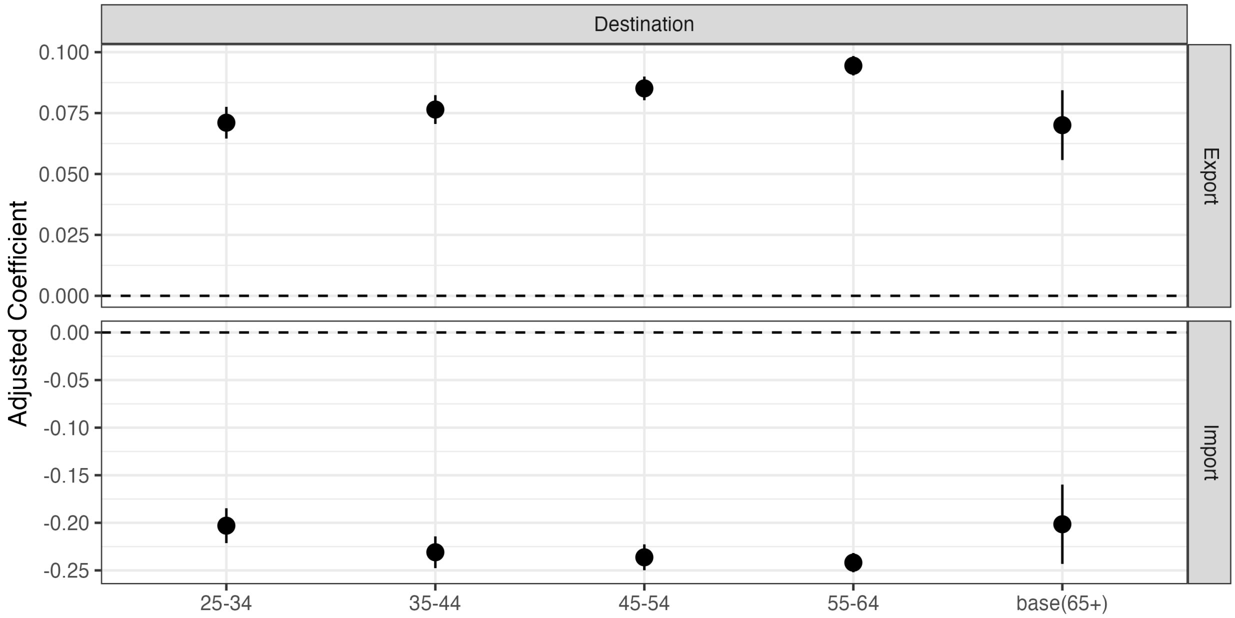


Figure 4: Destination trade shock

Notes: The figure plots the coefficients for the effect of destination trade shock (export and import) by age groups. All coefficient values are adjusted by the baseline age group (65+). The vertical line for each point denotes 95% confidence interval.

specifically those who fall within the age range between 45 and 64.

A potential explanation for the observed variations in response to the China shock might be rooted in the insights provided by Pierce and Schott (2020). Using mortality data in the U.S., Pierce and Schott (2020) found that relationship between China shock and mortality was strongest for middle-aged white males. They offered insights into this relationship by examining data on the makeup of the manufacturing workforce. According to the US Bureau of Labor Statistics, a significant portion of middle-aged white males held positions within the manufacturing sector, a sector particularly susceptible to the impacts of the China shock. Similar to the U.S., it is known that large proportion of middle-aged male in Korea holds positions within the manufacturing sector (Kim, 2021). If similar channel operated in Korea, we would observe a more sensitive migration response for households led by middle-age individual. Figure 4 seems to reveal such consistent patterns for the case of South Korea.

Another possible explanation for the observed phenomena could be related to accumulation of sector-specific skill for older workers. It is documented that certain skills exhibit age-dependent variations. Due to age-appreciating and age-depreciating skills of the workers, old workers have comparative advantage in certain sectors, whereas young workers have comparative advantage in certain sectors (Cai and Stoyanov, 2016). In this context, older workers could potentially be more vulnerable than younger workers to the trade shock if their age-related skills are specialized within a particular sector, and this specialization may impede their ability to adapt to the shock through transitioning to different sectors. On the contrary, young workers may be more flexible to change sectors in their early careers. Furthermore, Dix-Carneiro (2014) reports that older workers experience greater welfare losses compared to their younger counterparts when confronted with trade shock when estimating a structural dynamic equilibrium model of the Brazilian labor market. This is in part due to the fact that older workers have accumulated more sector-specific expertise than their younger counterparts. In this regard, the reason for the pronounced effect of mobility response in older head of households could be due to the fact that inter-sectoral transition cost is relatively higher for them. While younger head of households can switch their industry to adjust to the industry-specific trade shock, older head of households migrate as the cost of inter-sectoral transition outweighs the mobility cost.

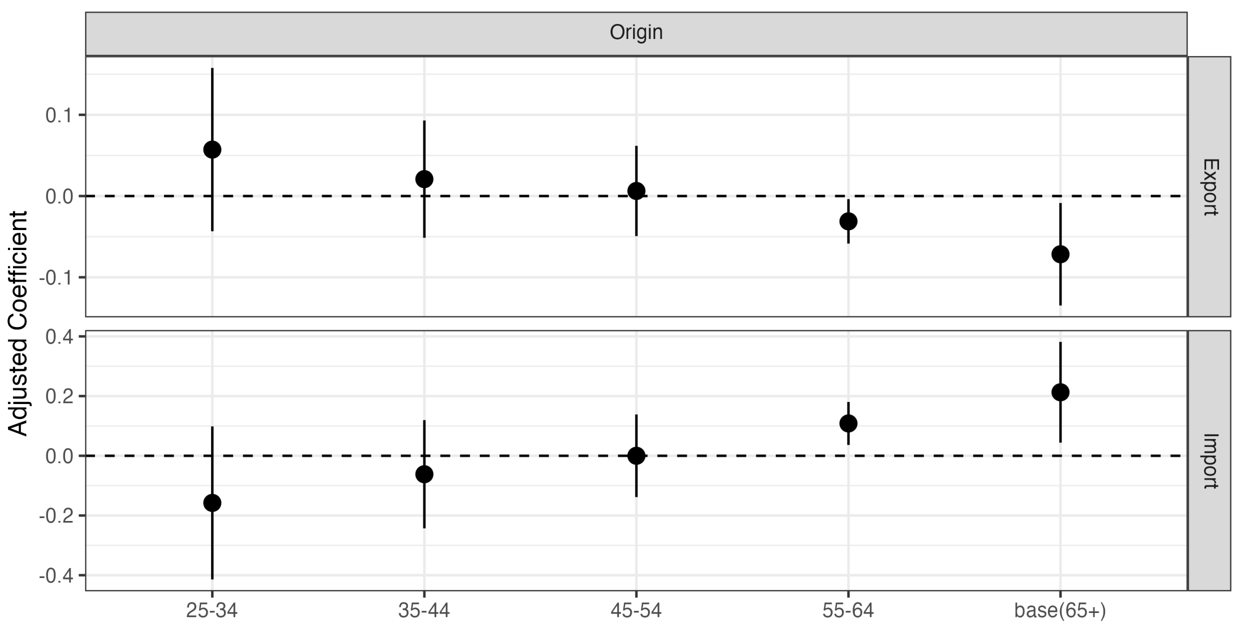


Figure 5: Origin trade shock

Notes: The figure plots the coefficients for the effect of origin trade shock (export and import) by age groups. All coefficient values are adjusted by the baseline age group (65+). The vertical line for each point denotes 95% confidence interval.

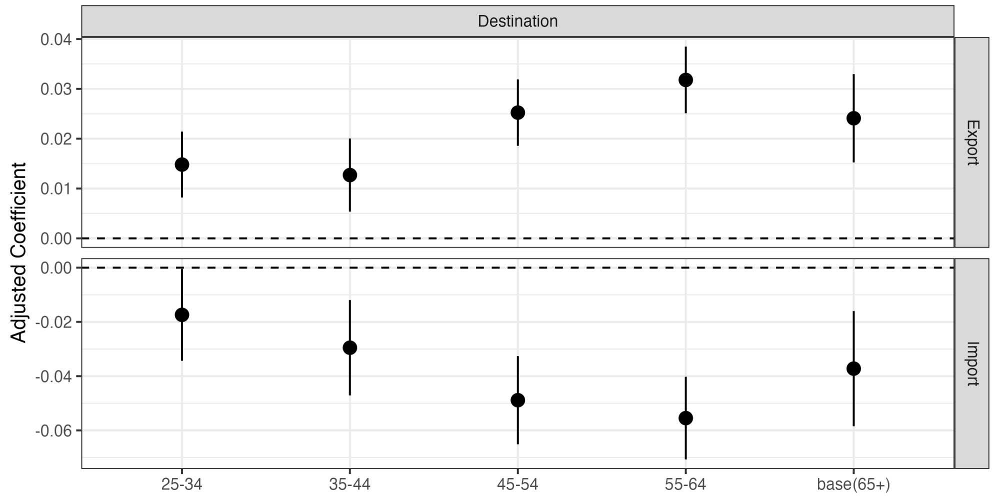
On the other hand, Figure 5 shows the results for the origin trade shock. Unlike the results in Figure 4, Figure 5 shows some striking patterns. While destination trade shock is fairly consistent for all age groups, origin trade shock's effects on migration differ greatly by age groups. In fact, it seems that effects of trade shock are monotonic by age groups. According to the figure, import penetration's effect on migration decreases as age of the head of household increases and export expansion's effect on migration increases. These intriguing patterns observed in the initial trade shocks could potentially explain the subtle effects of trade on migration, as highlighted in previous studies. It's evident that the effects of the origin trade shock are offset by varying impacts across different age groups. As this leads to only marginal effects in full sample, this might have been the potential factor leading to the past missing migration problem in trade literature.

5.2 Missing Migration Problem Revisited

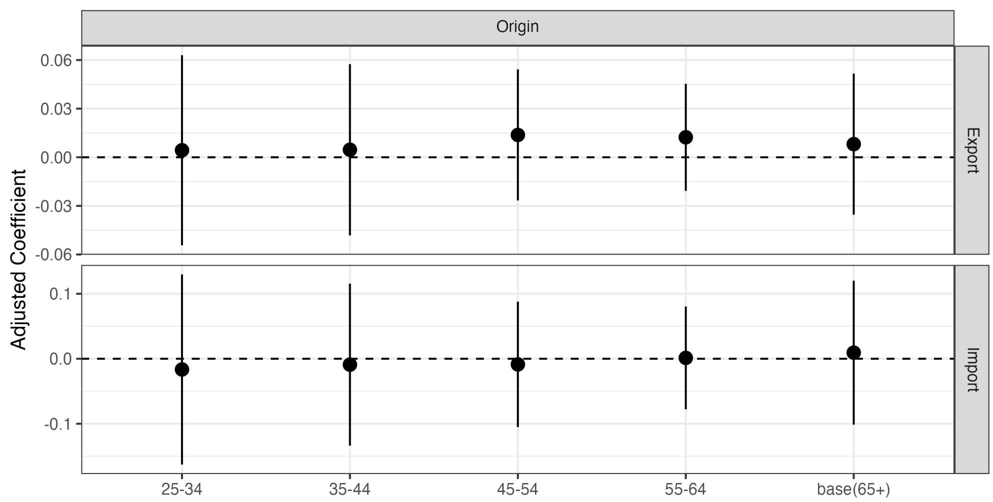
We found that the effects of origin trade shock were heterogeneous by age groups. Especially, we showed that the migration patterns for young head of households (25-34) were counter-intuitive to our understanding of the impact of the China trade shock on labor adjustment in the past literature. As these heterogeneous effects by age groups could be the main reason for the missing link between trade shock and migration, we further examine the potential source for such patterns.

To this end, we re-run our estimates for age groups by excluding commuting zone containing Seoul and only use job-related migration. This is because young individuals in South Korea typically exhibit a stronger preference for Seoul and its surrounding regions. For example, Kim et al. (2023) reported that young people in their 20s to 30s had significantly higher preference for Seoul metropolitan areas. Considering this fact, our baselines result could be a mixture of such preference for Seoul metropolitan areas by young head of households. We also focus on job-related migration to reduce noise in the migration patterns. When studying labor re-allocation through migration, it's essential to focus on job-related migrations. By narrowing our sample to this specific category, we can reduce the potential noise from migrations driven by reasons unrelated to employment. Table A.10 in the Appendix shows the results of the estimates. We again plot the coefficients of our results by age groups in Figure 6 to compare patterns between them.

Figure 6 shows that while the effects of destination trade shock are still consistent by age groups, the monotonic patterns in origin trade shock almost disappear and become null. This could partially explain the previous counter-intuitive findings arising from the origin trade shock. In metropolitan areas, such as the capital region, the movement of the population may be more influenced by external factors than by trade shocks. This is more plausible for country such as South Korea where there is a strong preference for the



(a) Destination trade shock



(b) Origin trade shock

Figure 6: Coefficient plots by age groups (Job-related migration, excluding Seoul metropolitan area)

Notes: Each figure plots the coefficients for the effect of destination trade shock and origin trade shock by age groups. All coefficient values are adjusted by the baseline age group (65+). The vertical line for each point denotes 95% confidence interval.

capital region. Also, we might remove additional noise from our analysis by only using job-related migrations which are more affected by trade shocks. Hence by removing two dimension of noises, the differential effects of age groups might have become homogeneous. On the other hand, the results for the destination trade shock are still consistent: Just like the results in previous section, the effect for destination trade shock has its peak in migrations of middle-aged head of households (45-64).

5.3 Gender Effect

We now examine if the impact of trade shock migration differs by gender. Table 3 shows the baseline results using migration of only male or female head of households.

Table 3: Baseline 2SLS Results – Male or Female head of household

	Dependent variable: log(migrate)					
	Male			Female		
	(1)	(2)	(3)	(4)	(5)	(6)
$(\Delta EX_{et} - \Delta IP_{et})$	0.011^{***}			0.003^{*}		
	(0.001)			(0.001)		
$(\Delta EX_{st} - \Delta IP_{st})$	0.006			0.006		
	(0.004)			(0.005)		
ΔIP_{et}		-0.024^{**}	-0.259^{***}		-0.009	-0.197^{***}
		(0.009)	(0.024)		(0.010)	(0.023)
ΔEX_{et}		0.014^{***}	0.094^{***}		0.004	0.071^{***}
		(0.003)	(0.008)		(0.003)	(0.008)
ΔIP_{st}		-0.088^{*}	-0.007		-0.112^{*}	-0.021^{*}
		(0.050)	(0.013)		(0.064)	(0.011)
ΔEX_{st}		0.031^{*}	0.002		0.039	0.006
		(0.019)	(0.005)		(0.024)	(0.005)
Controls	yes	yes	yes	yes	yes	yes
Period FEs	yes	yes	yes	yes	yes	yes
Region FEs	no	no	yes	no	no	yes
Num. of Period	2	2	2	2	2	2
Num. Obs.	92348	92348	92348	90321	90321	90321
R^2	0.749	0.738	0.719	0.761	0.745	0.763

Notes. All regressions in the table include same set of control variables mentioned in our empirical specification section. For the region fixed effects, we include origin commuting zone fixed effects and destination commuting zone fixed effects. Regression estimates are weighted by the initial population (by gender) in the year 2001. Standard errors are clustered at the district level; * p < 0.10, ** p < 0.05, *** p < 0.01.

We can see that destination trade shock from the baseline results still holds in both cases. However, it appears that the effects of trade shocks are more pronounced for migration of households with male representative. For example, columns (3) and (6) of Table 3 show that the estimates of the impact of destination import competition and export opportunity are -0.259 and 0.094 which are fairly larger than that of the female head of household (-0.197 and 0.071). In fact, the estimates for the trade shock on female head of households become insignificant in column (5).

This gendered result is well consistent with Autor et al. (2019) where they find that Chinese import competition reduced manufacturing labor demand, thereby exerting large negative impacts on men's relative employment and earnings in the US. The negative shock further affects marriage, fertility, premature mortality along with other multiple dimensions. If this channel had also operated in Korea, we would expect to observe more pronounced migration responses among male head of household than female head of household. Table 3 reveals such gender-specific migration reactions to the trade shock.

6 Conclusion

For the past few decades, much attention has been paid to identify the process of labor re-allocation of workers in response to trade shocks. However, various empirical studies have either found marginal or only limited evidence. This may be partially due to the limited data to precisely examine the effect of trade shock on internal migration. This paper addresses this largely unanswered question of the "missing migration problem" by providing novel data on migration. By leveraging detailed bilateral migration flows on the whole population of Korea, this paper provides new evidence that the labor reallocation in the case of trade shock occurs in terms of in-migration rather than out-migration. As most of the past studies on trade shock and migration focused on net population changes in regions (i.e. out-migration), we suggest that this may be the reason for the little mobility response results in past literature. In addition to introducing bilateral migration flows to separate the impact of trade shocks into destination and origin, we also decompose China shock into two channels: (i) import competition and (ii) export expansion. This enables us to precisely gauge the impact of trade shocks; without such refinement, our assessments of the broader effects of trade shocks on migration might be biased.

Overall, our findings are summarized as follows. The China trade shock affects in-migration but not out-migration of the regions. By separating the trade shock into export and import channel, we find that the export shock increases in-migration, whereas the import shock reduces in-migration. We find that effects of trade shock are most pronounced for middle-aged head of households between the age 45-64. Lastly, we also find that male head of households are more responsive to the effect of trade shock compared to female head of households. We anticipate that our findings will pave the way for more active investigations in this research arena.

References

Autor, David, David Dorn, and Gordon Hanson, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, October 2013, *103* (6), 2121–68.

_ **,** _ **, and** _ , "When work disappears: Manufacturing decline and the falling marriage market value of young men," *American Economic Review: Insights*, 2019, *1* (2), 161–178.

_ **,** _ **, and** _ , "On the Persistence of the China Shock," *Brookings Papers on Economic Activity*, 2021, pp. 381–447.

_ **,** _ **, and** _ , "Trading Places: Mobility Responses of Native and Foreign-Born Adults to the China Trade Shock," Working Paper 30904, National Bureau of Economic Research January 2023.

Balsvik, Ragnhild, Sissel Jensen, and Kjell G. Salvanes, "Made in China, sold in Norway: Local labor market effects of an import shock," *Journal of Public Economics*, 2015, *127*, 137–144. The Nordic Model.

Borusyak, Kirill, Rafael Dix-Carneiro, and Brian Kovak, "Understanding Migration Responses to Local Shocks," Working Paper, SSRN April 2022.

Cai, Jie and Andrey Stoyanov, "Population aging and comparative advantage," *Journal of International Economics*, 2016, *102*, 1–21.

Choi, Jaerim and Mingzhi Xu, "The labor market effects of the China Syndrome: Evidence from South Korean manufacturing," *World Economy*, 2020, *43* (11), 3039–3087.

Choi, Kyeongsoo and Jeongho Kim, "The Impact of Trade Expansion with China on Korean Labor Market (in Korean)," Report, KDI 2015.

Dauth, Wolfgang, Sebastian Findeisen, and Jens Suedekum, "The rise of the East and the Far East: German labor markets and trade integration," *Journal of the European Economic Association*, 2014, *12* (6), 1643–1675.

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

_ **and Brian K. Kovak**, "Margins of labor market adjustment to trade," *Journal of International Economics*, 2019, *117*, 125–142.

Donoso, Vicente, Víctor Martín, and Asier Minondo, "Do Differences in the Exposure to Chinese Imports Lead to Differences in Local Labour Market Outcomes? An Analysis for Spanish Provinces," *Regional Studies*, 2015, *49* (10), 1746–1764.

Facchini, Giovanni, Maggie Y. Liu, Anna Maria Mayda, and Minghai Zhou, "China's "Great Migration": The impact of the reduction in trade policy uncertainty," *Journal of International Economics*, 2019, *120*, 126–144.

Greenland, Andrew, John Lopresti, and Peter McHenry, "Import competition and internal migration," *Review of Economics and Statistics*, 2019, *101* (1), 44–59.

Hanlon, W. Walker, "Temporary Shocks and Persistent Effects in Urban Economies: Evidence from British Cities after the U.S. Civil War," *Review of Economics and Statistics*, 03 2017, *99* (1), 67–79.

Kim, Daeil, "The Impact of Chinese Economic Growth on the Korean Labor Market (in Korean)," Report, KDI 2006.

Kim, Hyoungchul, Seung Hoon Lee, and Ji Sub Park, "The Well-Being of Cities: Estimating Migration Attractiveness From Internal Migration Across Korean Cities," *Global Economic Review*, 2023, *52* (4), 313–325.

Kim, Youngbae, "S. Korea's workforce is aging faster than its competitors'," *Hankyoreh News*, 2021.

Koo, Kyong Hyun and Unjung Whang, "The Rise of China and the Rebound in Korea's Manufacturing Employment," Working Papers 18-7, Korea Institute for International Economic Policy December 2018.

Lee, Sewon and Hee-Yeon Lee, "Delimitation of City-Regions Based on the Method of Travel-to-Working Area and Analyzing Spatial Structure of City-Regions (in Korean)," *The Korea Spatial Planning Review*, 2015, *84*, 165–189.

Pierce, Justin R. and Peter K. Schott, "Trade Liberalization and Mortality: Evidence from US Counties," *American Economic Review: Insights*, March 2020, *2* (1), 47–64.

Twinam, Tate, "Trade competition and migration: Evidence from the quartz crisis," *Journal of International Economics*, 2022, *138*, 103653.

A Appendix

A.1 Data on Commuting Zone Characteristics

Most of the data for capturing regional characteristics and demographics (population, employment, gender ratio, etc) are from Korean Statistical Information Service (KOSIS) by Ministry of Statistics in South Korea. For most of our analysis, we use data from the year 2001 or earlier to capture the foundational characteristics of districts (or commuting zones) within South Korea prior to the onset of the China trade shock.

Table A.1 displays overall summary statistics of our data. Each panel features our main variables we use throughout our analysis. Panel A reports share of the cumulative (2001-2020) migrants moving across commuting zones out of total migration. We also show total number and the share of moving migrants by gender and age of the head of household (in millions). Panel B reports data related to employment in commuting zone such as average logarithmic value of total number of establishments, total employments and manufacturing employments for year 1999, 2000, 2001, 2010, and 2019. Finally, Panel C reports other control variables we use in our baseline analysis (e.g. population, share of female employment, etc).

Table A.1: Summary Statistics

Panel A: Migration Rates (Across commuting zone)		
	Cumulative, 2001-2020	**Total migrants**
Total migrants share	0.40	101.87
Male migrants share	0.35	53.41
Female migrants share	0.39	21.41
Age 25-34 migrants share	0.39	21.67
Age 35-44 migrants share	0.34	22.55
Age 45-54 migrants share	0.30	14.62
Age 55-64 migrants share	0.30	7.03
Age 65+ migrants share	0.34	4.09

Panel B: Employment by CZ					
	1999	**2000**	**2001**	**2010**	**2019**
Average Log(Employment)	10.32	10.36	10.37	10.57	10.87
Average Log(Manufacturing employment)	8.31	8.36	8.41	8.37	8.64

Panel C: Other control variables by district (D)		
	2001	**2010**
Average Log(Population) (D)	11.86	11.86
Average share of manufacturing employment (D)	0.19	0.12
Average share of female employment (D)	0.41	0.43

N of districts	226 districts
N of commuting zone	33 commuting zones
N of dest-origin pairs	$226 \times 226 - 226 = 50,850$
N of dest-origin pairs (Exclude same CZ)	46,664 obs

Notes: Values in the parentheses are standard errors. Unless noted, all of the variables are mean values for each district or commuting zone. All values are rounded to 2 decimal place. Values for "Total migrants" are in millions. Due to some NA values in the migration data, total number of migrants by group does not fully add up to the total number of migrants.

A.2 First-stage Results

Table A.2: First-stage results – equation (3)

	$(\Delta EX_{st} - \Delta IP_{st})$	$(\Delta EX_{et} - \Delta IP_{et})$
$(\Delta EX_{st} - \Delta IP_{st})$ IV	0.321^{***}	0.014^{***}
	(0.053)	(0.002)
$(\Delta EX_{et} - \Delta IP_{et})$ IV	0.007^{***}	0.417^{***}
	(0.002)	(0.002)
Controls	yes	yes
Period FEs	yes	yes
Num. of Period	2	2
Num. Obs.	93328	93328
R^2	0.347	0.328
F-statistic	32.7	33,606.3

Notes. All regressions in the table include same set of control variables mentioned in our empirical specification section. Regression estimates are weighted by the initial population in the year 2001. Standard errors are clustered at the district level; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: First-stage results – equation (4)

	ΔIP_{st}	ΔEX_{st}	ΔIP_{et}	ΔEX_{et}
ΔIP_{st} IV	0.240^{***}	0.273^{**}	0.014^{***}	-0.001
	(0.067)	(0.118)	(0.002)	(0.005)
ΔEX_{st} IV	0.095^{***}	0.374^{***}	-0.006^{***}	0.004^{**}
	(0.022)	(0.049)	(0.001)	(0.002)
ΔIP_{et} IV	0.002	-0.000	0.390^{***}	0.213^{***}
	(0.003)	(0.008)	(0.001)	(0.003)
ΔEX_{et} IV	-0.001	0.002	0.044^{***}	0.436^{***}
	(0.001)	(0.003)	(0.001)	(0.002)
Controls	yes	yes	yes	yes
Period FEs	yes	yes	yes	yes
Num. of Period	2	2	2	2
Num. Obs.	93328	93328	93328	93328
R^2	0.708	0.691	0.657	0.659
F-statistic	52.9	49.8	114,341.6	88,745.0

Notes. All regressions in the table include same set of control variables mentioned in our empirical specification section. Regression estimates are weighted by the initial population in the year 2001. Standard errors are clustered at the district level; * p < 0.10, ** p < 0.05, *** p < 0.01.

A.3 Robustness Checks

Table A.4: Baseline 2SLS Results – Age 25-64

	Dependent variable: log(migrate)		
	(1)	(2)	(3)
$(\Delta EX_{et} - \Delta IP_{et})$	0.013^{***}		
	(0.001)		
$(\Delta EX_{st} - \Delta IP_{st})$	0.010^{**}		
	(0.005)		
ΔIP_{et}		-0.031^{***}	-0.273^{***}
		(0.009)	(0.025)
ΔEX_{et}		0.016^{***}	0.100^{***}
		(0.002)	(0.009)
ΔIP_{st}		-0.115^{*}	-0.006
		(0.066)	(0.013)
ΔEX_{st}		0.043^{*}	0.005
		(0.024)	(0.006)
Controls	yes	yes	yes
Period FEs	yes	yes	yes
Region FEs	no	no	yes
Num. of Period	2	2	2
Num. Obs.	92632	92632	92632
R^2	0.768	0.750	0.722

Notes. All regressions in the table include same set of control variables mentioned in our empirical specification section. For the region fixed effects, we include origin commuting zone fixed effects and destination commuting zone fixed effects. Regression estimates are weighted by the initial population (25-64) in the year 2001. Standard errors are clustered at the district level; * p < 0.10, ** p < 0.05, *** p < 0.01.

Table A.5: Baseline 2SLS Results – One period

	Dependent variable: log(migrate)	
	(1)	(2)
$(\Delta EX_{et} - \Delta IP_{et})$	0.022^{***}	
	(0.002)	
$(\Delta EX_{st} - \Delta IP_{st})$	0.009	
	(0.010)	
ΔIP_{et}		-0.030^{***}
		(0.003)
ΔEX_{et}		0.018^{***}
		(0.001)
ΔIP_{st}		-0.021^{*}
		(0.012)
ΔEX_{st}		0.010
		(0.010)
Controls	yes	yes
Period FEs	no	no
Region FEs	no	no
Num. of Period	1	1
Num. Obs.	46644	46644
R^2	0.761	0.762

Notes. All regressions in the table include same set of control variables mentioned in our empirical specification section. For the region fixed effects, we include origin commuting zone fixed effects and destination commuting zone fixed effects. Regression estimates are weighted by the initial population in the year 2001. Standard errors are clustered at the district level; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Baseline 2SLS Results – Job-related migration

	Dependent variable: log(migrate)		
	(1)	(2)	(3)
$(\Delta EX_{et} - \Delta IP_{et})$	0.013^{***}		
	(0.001)		
$(\Delta EX_{st} - \Delta IP_{st})$	0.004		
	(0.004)		
ΔIP_{et}		-0.039^{***}	-0.213^{***}
		(0.009)	(0.024)
ΔEX_{et}		0.018^{***}	0.075^{***}
		(0.002)	(0.008)
ΔIP_{st}		-0.104^{*}	-0.071^{**}
		(0.055)	(0.035)
ΔEX_{st}		0.035^{*}	0.025
		(0.020)	(0.015)
Controls	yes	yes	yes
Period FEs	yes	yes	yes
Region FEs	no	no	yes
Num. of Period	2	2	2
Num. Obs.	92284	92284	92284
R^2	0.731	0.717	0.728

Notes. All regressions in the table include same set of control variables mentioned in our empirical specification section. For the region fixed effects, we include origin commuting zone fixed effects and destination commuting zone fixed effects. Regression estimates are weighted by the initial population in the year 2001. Standard errors are clustered at the district level; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: Baseline 2SLS Results – Excluding Seoul and Jeju-do

	Dependent variable: log(migrate)		
	(1)	(2)	(3)
$(\Delta EX_{et} - \Delta IP_{et})$	0.013^{***}		
	(0.001)		
$(\Delta EX_{st} - \Delta IP_{st})$	0.004		
	(0.006)		
ΔIP_{et}		-0.048^{***}	-0.096^{***}
		(0.010)	(0.016)
ΔEX_{et}		0.019^{***}	0.034^{***}
		(0.003)	(0.006)
ΔIP_{st}		-0.038	-0.041^{**}
		(0.050)	(0.020)
ΔEX_{st}		0.015	0.014
		(0.019)	(0.009)
Controls	yes	yes	yes
Period FEs	yes	yes	yes
Region FEs	no	no	yes
Num. of Period	2	2	2
Num. Obs.	74296	74296	74296
R^2	0.781	0.777	0.816

Notes. All regressions in the table include same set of control variables mentioned in our empirical specification section. For the region fixed effects, we include origin commuting zone fixed effects and destination commuting zone fixed effects. Regression estimates are weighted by the initial population in the year 2001. Standard errors are clustered at the district level; * p < 0.10, ** p < 0.05, *** p < 0.01.

Table A.8: Baseline 2SLS Results – Pre-exposure (1996-2000) migration

	Dependent variable: log(migrate)	
	(1)	(2)
$(\Delta EX_{et} - \Delta IP_{et})$	-0.018^{***}	
	(0.003)	
$(\Delta EX_{st} - \Delta IP_{st})$	-0.027^{**}	
	(0.013)	
ΔIP_{et}		-0.034^{*}
		(0.019)
ΔEX_{et}		-0.001
		(0.008)
ΔIP_{st}		0.155
		(0.243)
ΔEX_{st}		-0.074
		(0.100)
Controls	yes	yes
Period FEs	no	no
Region FEs	no	no
Num. of Period	1	1
Num. Obs.	46353	46353
R^2	0.742	0.691

Notes. All regressions in the table include same set of control variables mentioned in our empirical specification section. For the region fixed effects, we include origin commuting zone fixed effects and destination commuting zone fixed effects. Regression estimates are weighted by the initial population in the year 2001. Standard errors are clustered at the district level; * p < 0.10, ** p < 0.05, *** p < 0.01.

A.4 Regression Tables for Age Effect

Table A.9: Baseline 2SLS Results – Age (base is 65+)

	Dependent variable: log(migrate)	
	(1)	(2)
ΔIP_{st}	0.172^{***}	0.213^{**}
	(0.065)	(0.086)
$\Delta IP_{st} \times 1\{age = 25 \sim 34\}$	-0.375^{***}	-0.371^{***}
	(0.141)	(0.131)
$\Delta IP_{st} \times 1\{age = 35 \sim 44\}$	-0.275^{***}	-0.275^{***}
	(0.101)	(0.092)
$\Delta IP_{st} \times 1\{age = 45 \sim 54\}$	-0.214^{***}	-0.213^{***}
	(0.076)	(0.071)
$\Delta IP_{st} \times 1\{age = 55 \sim 64\}$	-0.107^{***}	-0.104^{***}
	(0.038)	(0.037)
ΔIP_{et}	-0.013	-0.202^{***}
	(0.013)	(0.021)
$\Delta IP_{et} \times 1\{age = 25 \sim 34\}$	-0.009	-0.002
	(0.010)	(0.009)
$\Delta IP_{et} \times 1\{age = 35 \sim 44\}$	-0.036^{***}	-0.029^{***}
	(0.009)	(0.008)
$\Delta IP_{et} \times 1\{age = 45 \sim 54\}$	-0.041^{***}	-0.035^{***}
	(0.007)	(0.007)
$\Delta IP_{et} \times 1\{age = 55 \sim 64\}$	-0.045^{***}	-0.040^{***}
	(0.005)	(0.005)
ΔEX_{st}	-0.059^{**}	-0.072^{**}
	(0.026)	(0.032)
$\Delta EX_{st} \times 1\{age = 25 \sim 34\}$	0.131^{**}	0.129^{**}
	(0.055)	(0.051)
$\Delta EX_{st} \times 1\{age = 35 \sim 44\}$	0.093^{**}	0.092^{**}
	(0.040)	(0.037)
$\Delta EX_{st} \times 1\{age = 45 \sim 54\}$	0.080^{**}	0.078^{***}
	(0.031)	(0.028)
$\Delta EX_{st} \times 1\{age = 55 \sim 64\}$	0.042^{***}	0.041^{***}
	(0.015)	(0.014)
ΔEX_{et}	0.005	0.070^{***}
	(0.004)	(0.007)
$\Delta EX_{et} \times 1\{age = 25 \sim 34\}$	0.004	0.001
	(0.004)	(0.003)
$\Delta EX_{et} \times 1\{age = 35 \sim 44\}$	0.009^{***}	0.006^{**}
	(0.003)	(0.003)
$\Delta EX_{et} \times 1\{age = 45 \sim 54\}$	0.018^{***}	0.015^{***}
	(0.003)	(0.002)
$\Delta EX_{et} \times 1\{age = 55 \sim 64\}$	0.027^{***}	0.024^{***}
	(0.002)	(0.002)
Controls	yes	yes
Period $\times$ Age FEs	yes	yes
Region FEs	no	yes
Num. of Period	2	2
Num. Obs.	422943	422943
R^2	0.719	0.711

Notes: All regressions in the table include same set of control variables mentioned in our empirical specification section. For the region fixed effects, we include origin commuting zone fixed effects and destination commuting zone fixed effects. Regression estimates are weighted by the initial population (by age) in the year 2001. Standard errors are clustered by districts.

Table A.10: Baseline 2SLS Results – Age (base is 65+), Job-related migration, excluding Seoul metropolitan area

	Dependent variable: log(migrate)	
	(1)	(2)
ΔIP_{st}	-0.114	0.009
	(0.184)	(0.056)
$\Delta IP_{st} \times 1\{age = 25 \sim 34\}$	-0.023	-0.026
	(0.088)	(0.075)
$\Delta IP_{st} \times 1\{age = 35 \sim 44\}$	0.002	-0.018
	(0.072)	(0.064)
$\Delta IP_{st} \times 1\{age = 45 \sim 54\}$	0.008	-0.018
	(0.055)	(0.049)
$\Delta IP_{st} \times 1\{age = 55 \sim 64\}$	0.016	-0.008
	(0.054)	(0.040)
ΔIP_{et}	-0.076^{***}	-0.037^{***}
	(0.011)	(0.011)
$\Delta IP_{et} \times 1\{age = 25 \sim 34\}$	0.019^{*}	0.020^{**}
	(0.011)	(0.009)
$\Delta IP_{et} \times 1\{age = 35 \sim 44\}$	0.006	0.008
	(0.011)	(0.009)
$\Delta IP_{et} \times 1\{age = 45 \sim 54\}$	-0.012	-0.012
	(0.010)	(0.008)
$\Delta IP_{et} \times 1\{age = 55 \sim 64\}$	-0.015^{*}	-0.018^{**}
	(0.009)	(0.008)
ΔEX_{st}	0.053	0.008
	(0.061)	(0.022)
$\Delta EX_{st} \times 1\{age = 25 \sim 34\}$	-0.003	-0.004
	(0.038)	(0.030)
$\Delta EX_{st} \times 1\{age = 35 \sim 44\}$	-0.009	-0.003
	(0.034)	(0.027)
$\Delta EX_{st} \times 1\{age = 45 \sim 54\}$	-0.003	0.006
	(0.024)	(0.021)
$\Delta EX_{st} \times 1\{age = 55 \sim 64\}$	-0.006	0.004
	(0.021)	(0.017)
ΔEX_{et}	0.043^{***}	0.024^{***}
	(0.005)	(0.005)
$\Delta EX_{et} \times 1\{age = 25 \sim 34\}$	-0.011^{***}	-0.009^{***}
	(0.004)	(0.003)
$\Delta EX_{et} \times 1\{age = 35 \sim 44\}$	-0.012^{***}	-0.011^{***}
	(0.004)	(0.004)
$\Delta EX_{et} \times 1\{age = 45 \sim 54\}$	-0.000	0.001
	(0.004)	(0.003)
$\Delta EX_{et} \times 1\{age = 55 \sim 64\}$	0.004	0.008^{**}
	(0.004)	(0.003)
Controls	yes	yes
Period $\times$ Age FEs	yes	yes
Region FEs	no	yes
Num. of Period	2	2
Num. Obs.	201573	201573
R^2	0.610	0.737

Notes. All regressions in the table include same set of control variables mentioned in our empirical specification section. For the region fixed effects, we include origin commuting zone fixed effects and destination commuting zone fixed effects. Regression estimates are weighted by the initial population (by age) in the year 2001. Standard errors are clustered at the district level; * p < 0.10, ** p < 0.05, *** p < 0.01.

B Methodology for Crosswalking HS Product Codes to KSIC Industry Codes

The process of crosswalking the HS product codes to the KSIC industry codes is critical for ensuring consistent and comparable data analytics. The detailed methodology followed for this crosswalk is as follows:

1. **Data Sources**: The base data consists of three primary datasets:
   - South Korea-China Import and Export Data
   - Japan-China Import and Export Data
   - Global-China Import and Export Data

   These datasets span three distinct years: 2001, 2010, and 2019.

2. **Variations in HS Codes**: The HS (Harmonized System) codes differ across the three datasets:
   - 2001 dataset uses the HS96 code system
   - 2010 dataset employs the HS07 code system
   - 2019 dataset adheres to the HS17 code system

3. **Crosswalking Framework**: The primary linkage between the HS codes and KSIC (Korea Standard Industry Classification) codes is established through the CPC (Central Product Classification) as an intermediary. The process of linkage was executed in the following manner:
   - HS1996 was mapped directly to HS2007
   - HS2007 was mapped to CPC version 2, which was subsequently linked to CPC version 2.1. This was finally crosswalked to the KSIC 10 with a 5-digit specificity.
   - HS2017 followed a more direct route, connecting to CPC version 2.1 and then to KSIC 10 with 5-digit granularity.

4. **KSIC Variations**: The National Business Survey (전국사업체조사) data contains three iterations of the KSIC: the 8th, 9th, and 10th versions. For the purpose of this study, all versions were standardized to the 10th version to ensure uniformity.

5. **Handling n:1 Relationships**: In instances where multiple HS codes mapped to a single KSIC code, the aggregation was straightforward as the data from the multiple HS codes was summed to align with the single KSIC code, ensuring there were no discrepancies.

6. **Addressing 1:n Relationships**: Challenges arose when a single HS code mapped to multiple KSIC codes. In such scenarios, the volume associated with the singular HS code was evenly distributed among the resultant KSIC codes. For example, if an HS code designated as “100" was associated with $100, and this code was cross-referenced with two KSIC codes (001 and 002), each KSIC code would be allocated $50.

7. **KSIC Code Handling**: Similar strategies were employed for KSIC codes as well. Whenever a code mapped to multiple resultant codes, the employment numbers were distributed evenly. This resulted in some figures not being whole numbers due to the division.

By following this detailed methodology, we ensured that the crosswalked data maintained its integrity and provided a consistent framework for analysis.

C Methodology for Concordance of Districts

The comprehensive methodology applied for ensuring a consistent concordance of districts in the data spanning from 1996 to 2020 is delineated below:

1. **Base Units for Concordance**: The foundational unit for concordance in this data set comprises 226 districts. These districts were used as the key reference points for integrating data over the period from 1996 to 2020.

2. **District Mergers**: During the observed time frame, certain districts underwent administrative mergers. In these instances, data from previously separate districts was consolidated. For analytical consistency, we retroactively combined data from such merging districts, treating them as a single unit, even for periods prior to the official merger. This ensured uniformity across the entire time span.

3. **District Splits**: Conversely, there were occasions when a single administrative district was bifurcated into multiple distinct entities. In such scenarios, data from these newly-formed districts was aggregated back into the original singular unit for the entirety of the observational period. This approach was adopted to preserve a consistent analytical framework that allows for clear and unambiguous comparative analyses across the years.

Incorporating these strategies has been instrumental in maintaining the integrity and clarity of the regional data, thus providing a robust foundation for meaningful analysis.