Submitted to *Econometrica*

SUPPLEMENTARY APPENDIX TO "RURAL PENSIONS, LABOR REALLOCATION, AND AGGREGATE INCOME: AN EMPIRICAL AND QUANTITATIVE ANALYSIS OF CHINA"

QINGEN GAI
Antai College of Economics & Management, Shanghai Jiao Tong University

NAIJIA GUO
HKU Business School, The University of Hong Kong

BINGJING LI
HKU Business School, The University of Hong Kong

QINGHUA SHI
Antai College of Economics & Management, Shanghai Jiao Tong University

XIAODONG ZHU
HKU Business School, The University of Hong Kong

Qingen Gai: gaiqingen@sjtu.edu.cn
Naijia Guo: njguo@hku.hk
Bingjing Li: bingjing@hku.hk
Qinghua Shi: shq@sjtu.edu.cn
Xiaodong Zhu: xdzhu@hku.hk

APPENDIX A: DATA APPENDIX

A.1. *Hukou Index*

We extend the prefecture-level *hukou* policy liberalization index constructed by Fan (2019) to 2013. Specifically, we search for and review all *hukou*-related official news articles as well as laws and regulations at the prefecture level from Peking University's Law Information Database and Baidu. Following the narrative approach by Fan (2019), we rate each document describing *hukou* policies on a score of 0 to 6, with 0 being the most stringent and 6 being the most liberal.[1] The average policy liberalization index increased from 2.04 in 2003 to 3.31 in 2010, and to 3.74 in 2013. This Hukou Index is a measure of a migrants job stability and the prospect of long-term settlement at a particular destination city in a particular year. Using the 2000 and 2010 population censuses, as well as the 2005 population mini census, Fan (2019) finds that a one-point increase in the destination-based Hukou Index leads to a 19%–21% increase in the number of inward migrants in the destination prefectures. These findings highlight the correlation between the Hukou Index and migration patterns.

To construct an origin-based annual Hukou Index faced by potential out-migrants from different localities, we proceed as follows. First, for each prefecture, we use the 2000 Population Census to calculate the shares of out-migrants to different destination prefectures. Second, employing the predetermined migration shares as weights, we calculate the average of the *hukou* policy liberalization indices across different destination prefectures. This measure is negatively related to the migration barriers faced by potential out-migrants from different localities and is referred to as the Hukou Index in the paper. Lastly, we assign the prefecture-level Hukou Index measures to the associated villages.

[1]See the details of the rating criteria in the appendix of Fan (2019). In the data, for each prefecture-year observation, there is at most one document on the *hukou* policy. If such a document exists, the score of the document is the Hukou Index for the prefecture in a given year. If there is no new document introducing new *hukou* reforms, we adopt the measure from the preceding year.

In Table A.1, we employ the NFP data to examine the impact of *hukou* policy liberalization in destination prefectures, as indicated by an increase in the origin-based Hukou Index, on out-migration flows from different localities. We employ a continuous difference-in-differences analysis to assess this relationship. The dependent variable in Column (1) is our baseline measure of migration, which is equal to one if an individual worked more than 180 days out of town during the year, and zero otherwise. Our findings reveal a positive and statistically significant effect of the Hukou Index on migration. Specifically, a one-standard-deviation increase in the Hukou Index leads to a 2 percent increase in the probability of migration.[2] In Column (2), we additionally control for the effects of the NRPS and individual characteristics; the estimate for the Hukou Index remains robust. In Columns (3)–(4), we repeat the analysis while replacing the dependent variable by the number of working days spent out of town. The estimate in Column (3) indicates that a one-standard-deviation increase in the Hukou Index leads to a significant increase of 6.9 working days spent out of town.

A.2. *The NFP Data: Details*

Overview The inception of the National Fixed Points (NFP) Survey dates back to 1984–1985 when the Central Rural Policy Research Office (CRPRO) undertook a nationwide socio-economic survey to evaluate the impacts of reforms in rural areas. This extensive survey covered a sample of 37,422 farming households in 272 villages across 28 provinces. In 1986, the CRPRO made the decision to designate the surveyed villages from the 1984–1985 survey as long-term fixed points for continuous and comprehensive observation, with a planned duration of up to 50 years. The NFP system was established, and has served two main objectives since then: (i) to provide a comprehensive understanding of the grassroots situation in rural areas; and (ii) to evaluate rural policies. In its early years, the NFP survey was primarily conducted at the village and household levels. Since 2003, the NFP survey

[2] The standard deviation of the Hukou Index across individuals and years is 0.678.

has included an individual-level questionnaire as an integral part of the data collection process. The NFP survey has periodically updated and rotated the samples within villages, and has incorporated new villages into the survey to enhance its coverage and representativeness. In this study, we use the data spanning the period from 2003 to 2013.

The NFP data is credible for several reasons. First, it is centrally managed by the Chinese Ministry of Agricultural and Rural Affairs, which ensures the direct reporting of data. The administrative structure of the survey comprises four levels, encompassing 31 provincial supervisory departments, 68 county-level supervisory departments, over 2,000 county and village investigators, and over 1,000 village assistant investigators. Direct reporting to the central government effectively mitigates the risk of data manipulation or falsification by provincial or municipal governments. Second, since its introduction in 1986, the bookkeeping system has played a crucial role in ensuring the accuracy of the original data. More specifically, bookkeeping books are issued to farmers, allowing them to enter their information in a timely manner. The records in these books are then utilized and transferred to the corresponding record sheets and survey sheets. Third, assistant investigators, often village officials or farmers acquainted with the community, are employed and trained to improve the quality of the survey. They conduct regular checks, verify farmers' bookkeeping activities, and guide farmers who need assistance to maintain their accounts. They also collaborate with county-level investigators to compile the data in a timely way. Lastly, the system rewards farmers and investigators with excellent bookkeeping records, which encourages continued adherence to best practices.

Comparing the NFP and Other Datasets Relative to repeated cross-sectional data, such as the population census, the panel structure of the NFP better serves identification purposes. Another advantage of the NFP over the population census is that the NFP provides detailed information on individual income, whereas only the 2005 mini population census includes information on income. The NFP has a much more comprehensive sample coverage in both geographical and time dimensions than other longitudinal surveys, such as the

Longitudinal Survey on Rural-Urban Migration in China (RUMiC) and the China Family Panel Study (CFPS). It tracks both rural residents and migrants annually over an 11-year period that encompasses the introduction of the NRPS. In particular, given that the NFP is an origin-based survey, its attrition rate is much lower than the destination-based surveys of migrants such as the RUMiC.

One drawback of the NFP data is the limited information on migration destinations. We have information only on whether a migrant is within the home county, within the home province, or outside the home province. For the surveys after 2009, we know the destination provinces but not the destination cities. Hence, the population census data are more suitable for analyzing the spatial allocation of labor.

To evaluate the representativeness and quality of the NFP data, we compare the 2005 wave of the NFP survey with a randomly selected 10% sample from the 2005 China 1% Population Sampling Survey (mini census). As reported in Table A.2, the observations in the NFP data exhibit similar characteristics to those of individuals holding rural *hukou* in the mini census, particularly with regards to educational attainment, labor force participation, the proportion of elderly individuals, and the share of agricultural employment.[3] We also aggregate the NFP data to facilitate a comparison with the national-level data reported by the National Bureau of Statistics (NBS). Given that the NFP was established in 1986 and its sample selection was primarily based on rural development at that time, it is reasonable to expect potential disparities in the levels of the two data series. Nevertheless, the two data series demonstrate comparable trends throughout the sample period, as depicted in Figure A.1.

A.3. *Sample Attrition in the NFP*

The NFP employs a variety of methods to ensure the effective tracking of migrants. First, the surveys are strategically conducted after December each year, aligning with the Chinese

[3]We repeat the analysis by comparing the 2010 wave of the NFP survey with a randomly selected 1% sample from the 2010 Population Census. The findings are consistent (the results are available upon request).

Spring Festival when most migrants return to their hometowns. Second, the local investigators are acquainted with the surveyed households, with access to essential information and contact details. As telephones and cell phones have become increasingly prevalent in China, these investigators can maintain communication with migrants through various means. These measures and strategies play a crucial role in enabling successful follow-up surveys on rural households and their members, particularly migrants.

To gain a better understanding of sample attrition in the NFP and the underlying factors, we compute the attrition rates at both the individual and household levels. The NFP data assigns unique codes to households but not to individual household members. Hence, we use the household code to track households and we employ the household code in conjunction with the gender and age of individuals in 2003 as identifiers to track individuals. The yearly attrition rate is defined as:

$$\text{Raw attrition rate}_t = 1 - \frac{\#\text{ of these observations tracked at t+1}}{\#\text{ of observations at t.}}$$

To account for the possibility that respondents may reappear in subsequent periods even after being absent in the immediate follow-up, we adjust the measure of attrition rate according to:

$$\text{Adjusted attrition rate}_t = 1 - \frac{\#\text{ of these observations tracked in any period after t}}{\#\text{ of observations at t.}}$$

This alternative measure reflects the overall tracking situation for the entire sample period.

Panel A in Figure A.2 finds that between 2003 and 2012, the raw and adjusted attrition rates at the household level average 8.01% and 3.89%, respectively, while the raw and adjusted attrition rates at the individual level average 20.26% and 15.26%, respectively.[4] As revealed in Panel B, there is no discernible difference in attrition rates between

[4]Note that the lack of unique codes assigned to NFP household members may result in inaccuracies in identifying follow-up samples, leading to an overestimate of sample attrition.

migrants and stayers. To benchmark the sample attrition rates in the NFP, we make a comparison with the CFPS, which is widely adopted in other studies, such as Lagakos et al. (2020). The CFPS employs an internationally-accepted survey methodology that assigns a unique identification code to each observed household and individual. The raw and adjusted household-level attrition rates in the CFPS data are 17.02% and 12.51%, respectively, over the period of 2010 to 2018. Similarly, at the individual level, the corresponding rates are 23.19% (raw) and 15.66% (adjusted).

If sample attrition is selective, our empirical identification based on sectoral switchers could be biased. For example, if individuals with higher abilities are more likely to migrate and attrit from the sample, we may underestimate the income gap between the agricultural and non-agricultural sectors. To investigate this potential issue, we compare the characteristics in year $t-1$ between individuals who were successfully tracked in year t and individuals who were not tracked. The results are presented in Table A.3. We find that a higher proportion of individuals who were not tracked are female, have poor health, and are elderly. Furthermore, individuals in the labor force who were not tracked were more likely to have higher education levels, to be working in non-agricultural sectors, to be working fewer hours in the agricultural sector, and to have a greater likelihood of outmigration. Migrant workers who were successfully tracked work a similar number of days as migrant workers who were not tracked, but the former group earns more than the latter group. Hence, we conclude that the sample attrition in NFPS is likely to bias our estimates toward zero. However, it is important to note that the magnitude of this bias is likely to be small, considering the relatively low attrition rate discussed earlier.

A.4. *Construction of Key Variables: Details*

Sector of Employment and Migration The NFP provides the following information, which can be used to infer the sector of employment and earnings for each sector: (i) the number of working days in the within-town agricultural sector and the within-town non-agricultural sector; (ii) the number of working days out of town; (iii) net income from agricultural production at the household level; and (iv) income earned out of town at the individual level. Table A.4 shows that out-migration status and non-agricultural employment are highly correlated. Those who work more than 180 days out of town spend only 3.6% of their working days in agricultural production on average, and 91.8% of these workers report non-agriculture as their sector of employment. Meanwhile, among those who spend less than 180 working days out of town, the share of working days allocated to agricultural production is 78% (i.e, the weighted average of the statistics in Columns (1)–(2)), and the share of workers reporting non-agriculture as their sector of employment is only 19.4%. Column (2) shows that the majority of workers with out-of-town working days within the range $(0, 180]$ still report agriculture as their sector of employment.

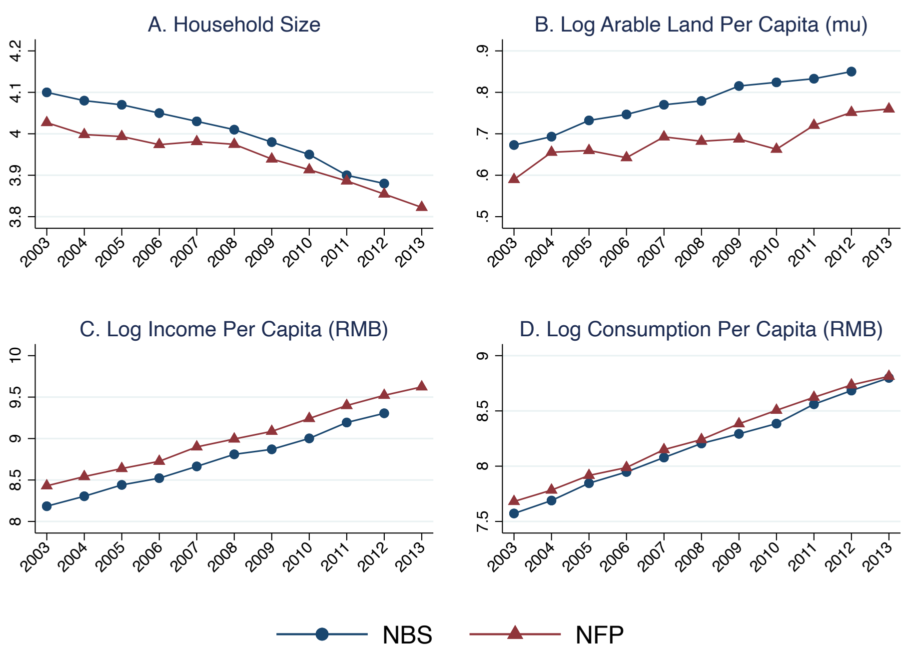
Based on these observations, the reduced-form analysis in this paper does not distinguish between sector choice and location choice. We loosely define sector of employment as follows: an individual is affiliated with the na sector if he works out of town for more than 180 days, and is affiliated with the a sector otherwise.[5] Panel A of Figure A.3 shows the distributions of working days allocated to within-town agriculture, within-town non-agriculture, and out of town for workers who are grouped in the a sector. We find that for workers in the a sector, 64.8% have zero working days in the within-town na sector and 90.8% spend zero working days out of town. Analogously, Panel B reveals that for workers in the na sector, 72.7% have zero working days in the within-town a sector and 95.1% have zero working days in the within-town na sector.

Deflating Nominal Earnings to Real Earnings We deflate all nominal earnings into 2003 Beijing prices using province-level spatial price deflators constructed by Brandt and Holz (2006). Specifically, for workers in agriculture, we deflate their daily earnings by the rural price index of the province in which their village is located. For out-of-town non-agricultural workers within their home province, we deflate daily earnings by the urban price index of the same province. For workers in the out-of-province non-agricultural sec-

[5]The NBS adopts a cutoff of 180 days to define migrant workers.

tor, their migration destination is unobserved during the period 2003–2008. To deflate their incomes, we proceed as follows. First, we use the 2000 Population Census to calculate the shares of out-migrants to different provinces for each prefecture. Second, we map the villages to prefectures, and based on the predetermined migration shares, construct the weighted average of urban price indices across different destination provinces for each village. The daily earnings of out-migrants are deflated by this weighted urban price index.

FIGURE A.1.—NBS versus NFP



Notes: The figure compares several variables aggregated from the NFP with those from the China Rural Statistical Yearbooks published by the National Bureau of Statistics (NBS). In 2013, the NBS stopped reporting the data on household size, arable land per capita, and income per capita in rural areas.

FIGURE A.2.—Attrition Rates

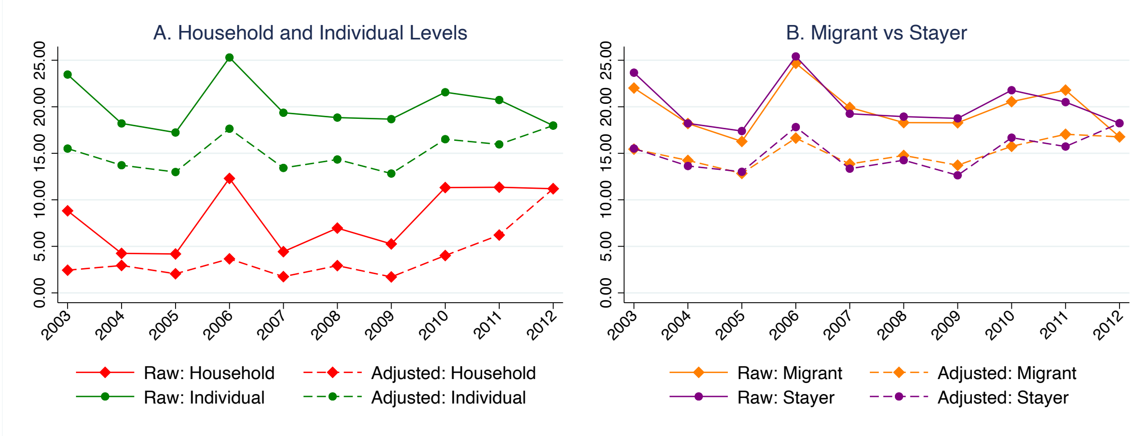
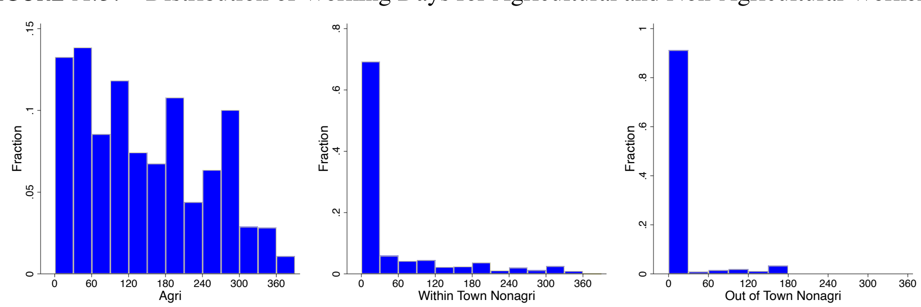
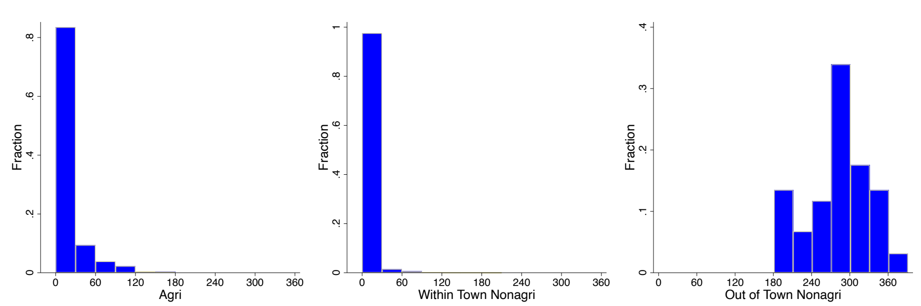


FIGURE A.3.—Distribution of Working Days for Agricultural and Non-Agricultural Workers

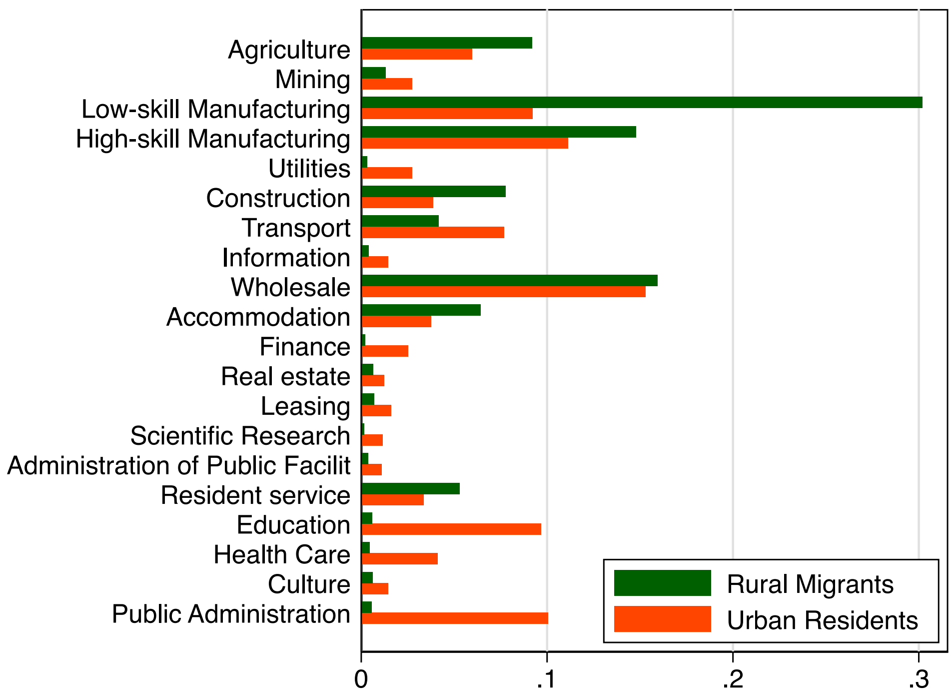


Panel A. Working days in different sectors for agricultural workers



Panel B. Working days in different sectors for non-agricultural workers

FIGURE A.4.—Sectoral Distribution across Rural Migrants and Urban Residents



Notes: We disaggregate the manufacturing sector into high-skill and low-skill manufacturing. Workers are defined as high-skilled if they have at least a college degree and as low-skilled otherwise. High-skill manufacturing represents the manufacturing industries that have a higher share of high-skilled workers than the median manufacturing industry.

TABLE A.1

ORIGIN-BASED HUKOU INDEX AND OUT-MIGRATION

Dep. Var.:	(1) NonAgri	(2) NonAgri	(3) Out-of-Town Working Days	(4) Out-of-Town Working Days
Hukou Index	0.0301	0.0228	10.6195	8.4248
	(0.0127)	(0.0115)	(3.7496)	(3.3844)
Elder60×NRPS		0.0419		10.8607
		(0.0075)		(2.2197)
NRPS		0.0116		3.6824
		(0.0100)		(3.0848)
Elder60		0.0232		7.2354
		(0.0026)		(0.7660)
Individual controls	N	Y	N	Y
Province× Year FE	Y	Y	Y	Y
Village FE	Y	Y	Y	Y
Observations	229,849	229,849	229,849	229,849
R-squared	0.1803	0.3611	0.1926	0.3850

Notes: Individual controls include age, age squared, years of education, a dummy for gender, and a dummy for poor health. Robust standard errors are clustered at the village×year level.

TABLE A.2

SUMMARY STATISTICS:
THE NFP AND THE 2005 CHINA 1% POPULATION SAMPLING SURVEY

	NFP	Census,2005	
		Rural Hukou	Urban Hukou
Age	36.9365	33.9327	36.8731
	(17.8483)	(20.4429)	(19.4439)
Female	0.4669	0.5022	0.4868
	(0.4989)	(0.5000)	(0.4998)
Years of Schooling	6.7794	6.5092	9.7278
	(3.0800)	(3.5588)	(4.2510)
Poor Health Status	0.0396	0.0311	0.0179
	(0.1950)	(0.1737)	(0.1326)
Share of Workers	0.7264	0.6130	0.6526
	(0.4458)	(0.4871)	(0.4761)
Share of Elders	0.0977	0.1134	0.1264
	(0.2969)	(0.3171)	(0.3323)
Share of Workers Working in Non-agriculture	0.4855	0.3739	0.9597
	(0.4998)	(0.4838)	(0.1966)
Share of Workers to Migrate	0.1665	0.1235	0.1983
	(0.3725)	(0.3290)	(0.3987)
Rural Migrant/Urban Resident's Annual Earnings (log)	8.7149	9.0762	9.3551
	(0.6399)	(0.5895)	(0.6173)
Share of Migrants Working in:			
Agriculture	0.0956	0.0939	0.0367
	(0.2941)	(0.2916)	(0.1881)
Industry	0.2555	0.4634	0.2655
	(0.4362)	(0.4987)	(0.4416)
Construction	0.1315	0.0772	0.0423
	(0.3380)	(0.2670)	(0.2013)
Service	0.5174	0.3654	0.6554
	(0.4997)	(0.4816)	(0.4752)

Notes: Standard deviation in parentheses.

TABLE A.3

CHARACTERISTICS OF TRACKED AND UNTRACKED INDIVIDUALS

	Tracked	Untracked	N	p-value
Age	37.4574	37.2361	874,695	0.000
	(19.5893)	(20.8165)		
Female	0.4739	0.4967	874,695	0.000
	(0.4993)	(0.5000)		
Poor Health	0.0464	0.0588	874,695	0.000
	(0.2104)	(0.2353)		
At School	0.1537	0.1458	874,695	0.000
	(0.3607)	(0.3529)		
In Labor Force	0.6618	0.6207	874,695	0.000
	(0.4731)	(0.4852)		
Elder	0.1263	0.1556	874,695	0.000
	(0.3321)	(0.3625)		
Among workers:				
Years of Schooling	7.1220	7.5704	561,648	0.000
	(3.0010)	(3.0755)		
Working in Non-agriculture	0.5150	0.5778	561,648	0.000
	(0.4998)	(0.4939)		
Migrant	0.2313	0.2569	561,648	0.000
	(0.4216)	(0.4369)		
Working days in Agriculture	82.7395	71.0459	561,648	0.000
	(105.7596)	(104.3250)		
Among migrants:				
Migrants Working Days	292.5246	294.5441	119,564	0.000
	(47.9879)	(46.7864)		
Migrants Real Earning(log)	9.3691	9.6307	119,564	0.000
	(0.7832)	(0.7947)		

Notes: Standard deviation in parentheses.

TABLE A.4

SUMMARY STATISTICS:
LABOR ALLOCATION AND SECTOR OF EMPLOYMENT BY OUT-OF-TOWN LABOR SUPPLY

	Agri Sector		Non-Agri Sector
Sample: Number of working days out of town	0 day	$(0, 180]$ days	> 180 days
	(1)	(2)	(3)
Total working days	205.7195	232.6875	302.2150
	(105.1715)	(75.0743)	(44.2010)
Share of working days in:			
Within-town Agri production	0.8162	0.4253	0.0356
	(0.3016)	(0.2267)	(0.0770)
Within-town NonAgri production	0.1838	0.0659	0.0049
	(0.3016)	(0.1429)	(0.0285)
Out-of-town	0.0000	0.5088	0.9595
	(0.0000)	(0.2347)	(0.0844)
(Self-reported) Non-agricultural sector	0.1742	0.3886	0.9180
	(0.3793)	(0.4875)	(0.2744)
ln Daily wage in Non-agricultural sector	0.0000	3.5347	3.4866
	(0.0000)	(0.7352)	(0.7255)
ln Daily wage in agricultural sector	2.9603	2.8840	2.9463
	(0.9691)	(0.9664)	(1.0350)
Number of observations	145,499	14,780	69,570

Notes: Standard deviation in parentheses.

APPENDIX B: ADDITIONAL DETAILS OF THE REDUCED-FORM EMPIRICAL ANALYSIS

In this section, we establish the theoretical framework that guides the empirical strategies employed in the reduced-form empirical analysis. The framework draws from a generalized Roy model, wherein rural workers migrate to urban non-agricultural sector if and only if the migration return exceeds the migration cost:

$$R + U_{na} - U_a > M(\boldsymbol{X}, \boldsymbol{Z}, T). \tag{B.1}$$

Here, U_{na} and U_a denote the unobserved individual productivity in na and a sectors, respectively; $R = \ln(w_{na}) - \ln(w_a)$ is the underlying real wage difference between the na and a sectors, i.e., APG; $M(\boldsymbol{X}, \boldsymbol{Z}, T)$ represents the effective migration costs faced by rural young workers, which depend on individual characteristics ($\boldsymbol{X}$), migration policies ($\boldsymbol{Z}$), and NRPS transfers to elder workers in the households (T).

This framework connects with the extant literature on migration and selection.[6] In addition, it aligns with the structural equation derived from the illustrative structural model outlined in Section 4.3 of the main text. In the structural model, we incorporate family structure, allow for interactions between elderly and young workers, and endogenize labor supply decisions. The model predicts that, due to an income effect, NRPS allows older workers to reduce their labor supply and allocate more time to home production. As a result, through a substitution effect, young workers in the same households can increase their market labor supply, which creates a stronger incentive for migration to the non-agricultural sector. Therefore, our structural model provides a micro-foundation for the effective migration cost $M(\boldsymbol{X}, \boldsymbol{Z}, T)$, which is a decreasing function of NRPS transfers T.

Our reduced-form empirical analysis aims at identifying the APG, R, and the migration cost, $M(\boldsymbol{X}, \boldsymbol{Z}, T)$. In the following subsections, we examine the potential sources of biases

[6] See Heckman and Honore (1990); Card (2001); Cornelissen et al. (2016); Pulido and Święcki (2018); Lagakos et al. (2020) among others.

when inferring the APG or migration cost from observed returns to migration. Furthermore, we detail our identification strategy and compare our estimates to those found in existing studies.

B.1. *Biases of Observed Returns to Migration*

If workers are homogeneous with identical labor supply and human capital in both sectors, then the underlying APG, R, is identical to the observed returns to migration. However, with endogenous labor supply and heterogeneous worker comparative advantage, observed returns to migration are often biased measures of the underlying APG. We will now discuss these potential biases in detail.

First, returns measured using annual earnings are biased estimates of APG because of endogenous labor supply, even if all workers are homogeneous and there is no sorting. Our theoretical model shows that young workers' labor supply is higher in non-agriculture than in agriculture as long as the migration cost is positive. This is consistent with the evidence on working days we presented for China in Table 1 in the main text and, more generally, consistent with the cross-country evidence on labor hours presented by Gollin et al. (2014). Our result suggests that the sectoral gap in labor supply itself may be a result of barriers to migration. Due to the gap in labor supply, returns measured using annual earnings overestimate the underlying APG and one should use hourly or daily earnings to avoid the bias.

Second, returns measured using daily earnings may still be biased due to heterogeneous human capital and sorting. Let $y_a(\boldsymbol{X},\boldsymbol{U}) = w_a h_a(\boldsymbol{X},\boldsymbol{U})$ and $y_{na}(\boldsymbol{X},\boldsymbol{U}) = w_{na} h_{na}(\boldsymbol{X},\boldsymbol{U})$ be the daily wages in the two sectors, respectively, where $\boldsymbol{U} = (U_a, U_{na})$. The observed log daily wage is given by

$$\ln y(\boldsymbol{X},\boldsymbol{U}) = \ln(w_a) + \mathbf{1}(j = na)R + \boldsymbol{X}\beta + U_a + \mathbf{1}(j = na)(U_{na} - U_a). \tag{B.2}$$

Let R_{OLS} be the observed difference in average log earnings of non-agricultural and agricultural workers, or *observed APG*, and $d = U_{na} - U_a$. We have

$$\begin{aligned} R_{\text{OLS}} &= E\left[\ln\left(y_{na}(\boldsymbol{X},\boldsymbol{U})\right)|d > M(\boldsymbol{X},\boldsymbol{Z},T) - R\right] - E\left[\ln\left(y_a(\boldsymbol{X},\boldsymbol{U})\right)|d \leq M(\boldsymbol{X},\boldsymbol{Z},T) - R\right] \\ &= R + E\left[U_{na}|d > M(\boldsymbol{X},\boldsymbol{Z},T) - R\right] - E\left[U_a|d \leq M(\boldsymbol{X},\boldsymbol{Z},T) - R\right] \end{aligned} \tag{B.3}$$

Due to heterogeneous innate abilities and sorting, the observed APG is generally different from the underlying APG. The last two terms in equation (B.3) show the selection bias or the effect of sorting on the deviation of the observed APG from the underlying APG. The sign of the bias is ambiguous, depending on both the joint distribution of (U_a, U_{na}) and the net migration cost faced by individuals, $M(\boldsymbol{X},\boldsymbol{Z},T) - R$. In the special case when (U_a, U_{na}) follows a bi-variate normal distribution, the selection bias has the following expression (see, e.g., Heckman and Honore, 1990):

$$R_{\text{OLS}} - R = \sigma_{na}\rho_{na,d}\frac{\phi(\frac{R - M(\boldsymbol{X},\boldsymbol{Z},T)}{\sigma_d})}{\Phi(\frac{R - M(\boldsymbol{X},\boldsymbol{Z},T)}{\sigma_d})} + \sigma_a\rho_{a,d}\frac{\phi(\frac{R - M(\boldsymbol{X},\boldsymbol{Z},T)}{\sigma_d})}{1 - \Phi(\frac{R - M(\boldsymbol{X},\boldsymbol{Z},T)}{\sigma_d})}, \tag{B.4}$$

where σ_a, σ_{na}, and σ_d are the standard deviations of U_a, U_{na}, and d, respectively; $\rho_{a,d}$ and $\rho_{na,d}$ are the correlations of d with U_a and U_{na}, respectively; and $\phi(\cdot)$ and $\Phi(\cdot)$ denote the probability density function and cumulative density function of a standard normal distribution, respectively.

B.2. *Empirical Strategies*

We now turn to the empirical methods for dealing with the selection bias problem. In the literature on APG, there are two commonly used methods in dealing with the selection bias problem. The first method assumes that the distribution of $(\exp(U_a), \exp(U_{na}))$ takes either a multivariate Fréchet or a multivariate log-normal distribution, and uses the moment matching method to estimate the distribution parameters, underlying APG, and migration costs. See, e.g., Lagakos and Waugh (2013), Pulido and Święcki (2018), Tombe and Zhu (2019), Hao et al. (2020), and Adamopoulos et al. (2022). As pointed out by Heckman and Honore (1990), however, the identification of Roy models is not robust to alternative distribution assumptions, and the estimation results are also not robust, depending critically on the functional form assumptions.

More recently, several authors have adopted a second method, using the observed labor returns of sector switchers in panel data as estimates of the APGs. See, e.g. Herrendorf and Schoellman (2018), Alvarez (2020), and Hamory et al. (2021). While this method does not rely on strong functional form assumptions, it is not clear what the observed labor returns of sector switchers really measure. Both Pulido and Święcki (2018) and Lagakos et al. (2020) provide examples showing that these returns may over- or under-estimate the underlying APG if the shocks that caused workers to switch sectors are correlated with individual comparative advantages. Schoellman (2020) argues heuristically that, if the shocks are independent of individual comparative advantage, the estimated return to migration for switchers is not the underlying APG but a measure of the average migration cost faced by switchers before the shocks hit.

Thus, neither of the two commonly used methods in the APG literature is ideal for dealing with the selection bias problem. We therefore draw on the methods employed in the extensive literature in labor economics for the identification and estimation of generalized Roy models (see, e.g., Card (2001), Eisenhauer et al. (2015), and Cornelissen et al. (2016)). These methods are discussed as follows:

Control Function Approach Using the terminology of this literature, the underlying APG is the average treatment effect (ATE) of migration: $R = E[\ln(y_{na}(\boldsymbol{X},\boldsymbol{U})) - \ln(y_a(\boldsymbol{X},\boldsymbol{U}))]$. To control for selection bias, the literature suggests using either field or natural experiments. For the case of China, we will use the gradual implementation of NRPS as a policy experiment and a control function approach to estimate the ATE or the underlying APG. Specifically, we estimate equation (B.2) controlling for proxies that capture the selection terms $E[U_a|\mathbf{1}(j=a),\boldsymbol{X},\boldsymbol{Z}]$ and $E[U_{na}|\mathbf{1}(j=na),\boldsymbol{X},\boldsymbol{Z}]$. To construct these proxies, one needs to infer the covariances between unobserved components, U_{na} and U_a, and the observable variables, $\mathbf{1}(j=na)$, $\boldsymbol{X}$, $\boldsymbol{Z}$, and T. To this end, we estimate the first stage selection equation, $\mathbf{1}(R-M(\boldsymbol{X},\boldsymbol{Z},T)+U_{na}-U_a>0)$, using NRPS transfers T as the excluded instrument.[7]

In the special case where U_{na} and U_a follow a joint-normal distribution, we implement the two-stage control function procedure as follows. In the first stage, the selection equation is estimated using a probit model. The second stage estimating equation is:

$$\begin{aligned}\ln y(\boldsymbol{X},\boldsymbol{U}) =&\gamma_0+\gamma_1\mathbf{1}(j=na)+\boldsymbol{X}\gamma_2+\gamma_3\mathbf{1}(j=na)\frac{\phi((\boldsymbol{X},\boldsymbol{Z},T)\hat{\zeta})}{\Phi((\boldsymbol{X},\boldsymbol{Z},T)\hat{\zeta})}\\&+\gamma_4\mathbf{1}(j=a)\frac{\phi((\boldsymbol{X},\boldsymbol{Z},T)\hat{\zeta})}{1-\Phi((\boldsymbol{X},\boldsymbol{Z},T)\hat{\zeta})}+\omega,\end{aligned}$$

where $\hat{\zeta}$ is a vector of estimates obtained from the first-stage probit estimation of the selection model.

We also implement the control function estimation in a less parametric way. Specifically, the second-stage estimating equation is:

$$\ln y(\boldsymbol{X},\boldsymbol{U})=\gamma_0+\gamma_1\mathbf{1}(j=na)+\boldsymbol{X}\gamma_2+\gamma_3\mathbf{1}(j=na)\times\hat{v}+\gamma_4\hat{v}+\omega,$$

where $\hat{v}$ represents the regression residual obtained from the first-stage estimation of a linear probability model. As shown in Card (2001), under the identification assumption that,

$$E[U_a|v]=\psi v \quad \text{and} \quad E[U_{na}-U_a|v]=\eta v,$$

the estimate of γ_1 is consistent for ATE.

[7] Specifically, T only enters the selection equation, but not the earning equations for $y_{na}(\boldsymbol{X},\boldsymbol{U})$ and $y_a(\boldsymbol{X},\boldsymbol{U})$.

IV Approach Using the policy experiment, we also estimate the local average treatment effect (LATE) that reveals the average return to migration for workers who are marginally affected by the policy (i.e., compliers). We implement the estimation of LATE using the IV strategy (Angrist and Pischke, 2009). Under the exclusion assumption of the policy instrument, the IV estimate of return to migration is also an estimate of the average migration cost faced by these marginal workers. This interpretation is formalized in the following proposition, which considers the introduction of NRPS that changes T from zero to a positive value:

Proposition *If the NRPS transfer T is independent of individual comparative advantage in the non-agricultural sector, $d = U_{na} - U_a$, then,*

$$\lim_{T \to 0} R_{\text{LATE}} = \frac{E\left[M(\boldsymbol{X}, \boldsymbol{Z}, 0) f\left(M(\boldsymbol{X}, \boldsymbol{Z}, 0) - R\right)\right]}{E\left[f\left(M(\boldsymbol{X}, \boldsymbol{Z}, 0) - R\right)\right]},$$

where $f(.)$ is the PDF of d.

Proof The LATE estimate captures the gains in the daily wage of sectoral switchers whose migration decisions are marginally affected by the NRPS:

$$R_{\text{LATE}} = E\left[\ln\left(\frac{w_{na}h_{na}}{w_a h_a}\right) | \ln\left(\frac{w_{na}h_{na}}{w_a h_a}\right) - M(\boldsymbol{X}, \boldsymbol{Z}, 0) < 0 < \ln\left(\frac{w_{na}h_{na}}{w_a h_a}\right) - M(\boldsymbol{X}, \boldsymbol{Z}, T)\right].$$

Let $d = U_{na} - U_a$, and $M = M(\boldsymbol{X}, \boldsymbol{Z}, 0)$. Then, $M(\boldsymbol{X}, \boldsymbol{Z}, T) = M - \Delta M$;

$$\ln\left(\frac{w_{na}h_{na}}{w_a h_a}\right) = \ln\left(\frac{w_{na}\exp\left(\boldsymbol{X}\beta + U_{na}\right)}{w_a \exp\left(\boldsymbol{X}\beta + U_a\right)}\right) = R + d;$$

$$R_{\text{LATE}} = R + E\left[d | M - R - \Delta M < d < M - R\right].$$

When T approaches to zero, ΔM also approaches to zero. Let $G(x) = \int_{-\infty}^{x} v f(v) dv$, and $p(M)$ be the PDF of M. The LATE estimator can be rewritten as:

$$R_{\text{LATE}} = R + \frac{\int \left(G(M-R) - G(M-R-\Delta M)\right) p(M) dM}{\int \left(F(M-R) - F(M-R-\Delta M)\right) p(M) dM}.$$

Note that,

$$\lim_{\Delta M \to 0} \frac{G(M-R) - G(M-R-\Delta M)}{\Delta M} = G'(M-R) = (M-R) f(M-R)$$

$$\lim_{\Delta M \to 0} \frac{F(M-R) - F(M-R-\Delta M)}{\Delta M} = F'(M-R) = f(M-R).$$

By L'Hôpital's rule,

$$\lim_{\Delta M \to 0} R_{\text{LATE}} = R + \frac{\int (M-R) f(M-R) p(M) dM}{\int f(M-R) p(M) dM} = \frac{E\left[M f(M-R)\right]}{E\left[f(M-R)\right]}.$$

B.3. *Individual Fixed Effects Estimation*

Table B.1 presents a comparison between the results obtained from the individual fixed effects (FE) model and those from the baseline OLS regression. Column (2) shows that our fixed effect (FE) estimate of the returns to migration is 37 log points, which is actually slightly larger than the OLS estimate reported in Column (1).

This result contrasts the findings in the existing studies, which have shown that the estimated returns to migration after controlling individual fixed effects are much smaller than cross-sectional OLS estimates based on panel data from other countries. (We review the fixed effect estimates in the extant literature in the following.) As is discussed in Schoellman (2020), if all the sector switches are driven by exogenous shocks to migration costs, the FE estimate measures the average migration cost faced by switchers before the shocks hit. Hence, one interpretation of the difference in the results between China and the other countries is that migration costs are much higher in China due to its rigid *hukou* system that explicitly restricts rural-to-urban migration.

Observational Migration Returns in Other Contexts: OLS v.s. FE Estimates Hamory et al. (2021) show that after controlling for individual fixed effects the estimated APG drops from 36 log points to 24 log points for Indonesia, and from 48 log points to 22 log points for Kenya.[8] Alvarez (2020) shows that controlling for individual fixed effects also leads to a large reduction in the estimated income gap between the manufacturing sector and the agricultural sector in Brazil, from 48 log points to 9 log points, as well as a large reduction in the estimated income gap between the service sector and the agricultural sector in Brazil, from 48 log points to 4 log points. Using the data from the US, Herrendorf and Schoellman (2018) find that the wage gains based on switchers are only 6%, much lower than the cross-sectional wage gap of 76%. These results suggest that the labor returns to migration are small in many countries.

Lagakos et al. (2020) use the China Family Panel Study (CFPS) data to estimate the return from switching sectors in China and find that the cross-sectional OLS estimate is significantly higher than the FE estimate. Their outcome measure is different from ours in two aspects. First, their measure is on an annual basis, while our baseline analysis focuses on daily wage earnings. Second, the gains of migration in Lagakos et al. (2020) are based on consumption per capita, which is probably a lower bound for income gains, because income elasticity of consumption is generally less than 1. In fact, when we use the real annual earning data from the CFPS, we obtain an OLS estimate of 1.20 and an FE estimate of 1.29. (The details are available upon request.)

---

[8] Relatedly, using annual income as the outcome measure, Pulido and Świecki (2018) find that the estimated sectoral income gap reduces from 54 log points to 33 log points in Indonesia when individual fixed effects are controlled for.

TABLE B.1

SECTOR OF EMPLOYMENT AND DAILY WAGE: OLS V.S. FE

Dep. Var.: ln Daily Wage	(1) OLS	(2) FE
NonAgri	0.3059	0.3669
	(0.0141)	(0.0156)
Individual controls	Y	Y
Province× Year FE	Y	Y
Village FE	Y	N
Individual FE	N	Y
Observations	229,849	229,847
R-squared	0.4174	0.6661

Notes: Individual controls include age, age squared, years of education, a dummy for gender, a dummy for poor health, a dummy indicating whether there is an elderly aged 60 or above residing in the household, and the share of months in year t that the NRPS has been in effect. Robust standard errors are clustered at the village×year level.

APPENDIX C: ADDITIONAL EMPIRICAL RESULTS

C.1. *Heterogeneity: IV Estimation*

With the heterogeneity of migration costs across different rural areas in China, the IV estimate captures the weighted average of baseline migration costs faced by the NRPS-induced sector switchers (see Appendix B.2). To further shed light on this interpretation, we divide villages into two groups depending on whether the average Hukou Index (which is negatively related to migration barriers) faced by out-migrants in 2009–2013 is above or below the median. Then we estimate the LATE for each group. Column (1) of Table C.1 reports the IV regression results. The IV estimates imply that among the compliers, working in the non-agricultural sector increases the daily wage by 100 log points in regions with high baseline migration costs (i.e., regions where the Hukou Index is below the median). The corresponding effect is 76 log points for regions with low baseline migration costs (i.e., regions where the Hukou Index is above the median).

C.2. *Robustness: Control Function Estimation*

We extend the control function model in several dimensions such that it is less dependent on functional form restrictions and requires a less stringent identification assumption. First, we estimate the first-stage selection equation by extending equation (2) with the interactions between the instrument and controls (except for the village and province-year fixed effects); doing this allows the NRPS to affect migration decisions in a more non-parametric way.[9] Using the residuals obtained from this augmented model ($\hat{\nu}_{ihjt}$), we estimate the following second-stage regression:

$$\ln y_{ihjt} = \gamma_1 NonAgri_{ihjt} + X_{ihjt}\gamma_2 + \eta NonAgri_{ihjt} \times \hat{\nu}_{ihjt} + \psi \hat{\nu}_{ihjt} + \varphi_j + \varphi_{pt} + u_{ihjt}.$$

[9]We use age group dummies and education group dummies to non-parametrically capture the effects of age and education on the migration decision.

Under the identification assumption that

$$E[U_{a,ihjt}|\nu_{ihjt}] = \psi\nu_{ihjt} \quad \text{and} \quad E[U_{na,ihjt} - U_{a,ihjt}|\nu_{ihjt}] = \eta\nu_{ihjt}, \tag{C.1}$$

the estimated coefficient γ_1 reflects the ATE. The regression result is reported in Column (2) of Table C.1. Second, in Column (3), we further include the quadratic term of the residual and its interaction with $NonAgri$. This specification relaxes the linearity assumption in (C.1) (Wooldridge, 2015). Third, as noted in Card (2001), in a general setting, changes in the instrumental variable may affect the entire mapping between unobserved abilities and the outcome of interest, which leads to a violation of assumption (C.1).[10] Following Card (2001), to address the problem, Column (4) extends the control function approach by adding an interaction term of the residual with Z and a three-way interaction with $NonAgri \times Z$. Across these extended models, the estimates of γ_1 remain stable, ranging from 28.9 to 29.5 log points.

C.3. *Mechanisms: Additional Results*

C.3.1. *Heterogeneity: The Effect of NRPS on Migration*

Table C.2 presents the heterogeneous effects of the NRPS on the sector of employment by gender and by the presence of young children in a household, which provides indirect evidence for the mechanisms associated with the demand for home production in our model. Columns (1)–(2) indicate that there is a more pronounced effect of $Elder60_{hjt} \times NRPS_{jt}$ for female workers. Columns (3)–(5) show that the responses of migration decision to

[10] In this case,

$$Cov(U_a, \nu|Z=1) \neq Cov(U_a, \nu|Z=0), \quad Cov(U_{na} - U_a, \nu|Z=1) \neq Cov(U_{na} - U_a, \nu|Z=0),$$

which violate assumption (C.1). Nevertheless, a simple extension of the control function can be implemented with the identification assumptions:

$$E[U_a|\nu] = \eta_0(1-Z)\nu + \eta_1 Z\nu \quad \text{and} \quad E[U_{na} - U_a|\nu] = \psi_0(1-Z)\nu + \psi_1 Z\nu.$$

$Elder60_{hjt} \times NRPS_{jt}$ is stronger for households with a child aged 15 or below, but only for female workers. These findings are consistent with our proposed mechanism: in the context of rural China, female workers engage more in home production, such as childcare. As the NRPS allows elders to reallocate time from farm work to non-farm home production such as taking care of their grandchildren, we should expect the effect of $Elder60_{hjt} \times NRPS_{jt}$ on migration to be larger for female workers.

Table C.3 explores the heterogeneous effects by education group and age group. Columns (1)–(3) find that the introduction of NRPS has a larger impact on migration among workers with 0–6 years of education (i.e., primary school education or below). Columns (4)–(5) reveal that the NRPS effect is stronger for the age group of 40–55. The findings suggest that compliers are more likely to be workers with low levels of education or of older age, who tend to face higher migration costs.[11]

Columns (1)–(3) of Table C.4 explore the effect of $Elder60_{hjt} \times NRPS_{jt}$ by the location of non-agricultural employment. The effect is statistically significant only when the non-agricultural employment is outside the county of the registered Hukou. We interpret these findings as supportive evidence for our mechanism. Specifically, sector-switching costs associated with home production increase with migration distance. Hence, if the NRPS alleviates these costs, we should expect non-agricultural employment in more distant locations to be a more relevant margin of adjustment.

The findings in Tables C.2 to C.4 also indicate that compliers are more likely to have larger baseline migration costs, e.g., female workers who are more involved in home production, rural workers with lower educational attainment, and individuals who are more inclined towards cross-county rather than within-county migration. This pattern may explain the large IV/LATE estimate in Column (3) of Table 3.

[11] We interpret the results in conjunction with the structural estimates in Table 5 of the revised manuscript: (i) Migration costs decline with years of education; and (ii) The life-cycle migration cost exhibits a hump shape, peaking at age 46, which indicates that the average migration costs are higher for the age group of 40–55 than the age group of 20–39.

Lastly, in Column (4), we allow the effect of the NRPS to vary by the age of the elderly household member. Relative to households without an elderly member, the probability of being employed in a non-agricultural sector increases by 0.5%, 5.1%, and 3.6% following the introduction of the NRPS, for workers from households with an elderly member in the age group of 55–59, 60–69, and 70 or above, respectively. Individuals aged 55–59 are not entitled to NRPS transfers, and hence the insignificant estimate of $Elder55 - 59 \times NRPS$ serves as a placebo test in support of our mechanism. More importantly, the effect is the most pronounced for households with an elderly member aged 60–69, for whom the labor supply channel should be the most relevant. In line with these findings, Table C.5 reveals that the impacts of the NRPS on elderly workers' labor supply are significant only among elderly individuals below the age of 70 or those in good health.

C.3.2. *Potential Omitted Channels*

The introduction of the NRPS may have a direct impact on potential earnings that differ between households with and without an elderly member. For example, if households with an elderly member increase their investment in fixed capital for production with the NRPS transfers, the exclusion restriction is violated, leading to a bias of our IV estimate. To alleviate such concerns, Table C.6 explores whether the introduction of the NRPS affects saving and investment behaviors differentially for households with different structures in year t. We find that households with and without an elderly member are not differentially impacted by the introduction of the NRPS along these dimensions: (i) savings rate ($= \frac{\text{NetIncome - Consumption}}{\text{NetIncome}}$, where $NetIncome$ is the total income net of tax and production costs); (ii) investment in fixed capital for production; (iii) amount of loans; (iv) changes in the size of farmland under operation; and (v) the use of intermediate inputs per unit of labor in agricultural production.

The NRPS cash transfers may also encourage out-migration among individuals from households that were previously constrained by limited access to credit. To investigate this alternative channel, we examine the heterogeneous impacts of the NRPS across households with varying levels of wealth in Column (1) of Table C.7. We divide households into two groups based on whether their deposits are above or below the median. Likewise, in Column (3), households are categorized into two groups based on their level of cash holdings. If liquidity/credit constraints serve as the primary channel through which the NRPS influences migration decisions, the effect should be stronger among households with lower levels of wealth or cash holdings. However, there are no discernible heterogeneous impacts of $Elder60 \times NRPS$ across different groups. Columns (2) and (4) replicate the analysis by dividing households into tertiles based on their deposits or cash holdings, respectively; the results are consistent.

C.4. *Human-Capital-Adjusted Daily Agricultural Earnings*

In the baseline analysis, we impute the individual-level agricultural daily earnings by apportioning the household-level agricultural income to each household member based on their respective number of working days in the agricultural sector. A concern is that this imputation approach may underestimate the agricultural earnings of individuals with a stronger absolute advantage in agriculture within the household, resulting in an upward bias for the estimated sectoral income gap if these individuals have a higher propensity of migrating to the non-agricultural sector. To address this issue, we consider an alternative measure of nominal daily agricultural earnings that accounts for the observed differences in human capital across household members. First, using the 2005 mini population census, we estimate a Mincer regression for workers in the agricultural sector, relating log agricultural monthly income to a dummy for females, years of education, age, and age squared. Second, using the Mincer estimates, we impute the individual-level human capital according to: $hc_i = \exp(-0.1177 \times Female + 0.0406 \times Educ + 0.0215 \times Age - 0.0002 \times Age^2)$. For individual i in household h, his daily nominal earnings is given by: $\frac{hc_i \times i\text{'s working days in } a}{\sum_{j \in h} hc_j \times j\text{'s working days in } a} \times \frac{\text{Household } h\text{'s value-added from } a}{i\text{'s working days in } a}$. This nominal income is then converted to real earnings based on the same procedure described in Appendix A.4.

Table C.8 employs the human-capital-adjusted measures and repeats the analysis in Section 3. Across the three different specifications — OLS, IV, and CF — the estimates resemble those in our baseline analysis.

TABLE C.1

SECTOR OF EMPLOYMENT AND DAILY WAGE: ADDITIONAL RESULTS

Dep Var: ln Daily Wage	(1) IV	(2) CF	(3) CF	(4) CF
Hukou Index: below median × NonAgri	1.0063			
	(0.3890)			
Hukou Index: above median × NonAgri	0.7640			
	(0.3568)			
NonAgri		0.2954	0.2945	0.2885
		(0.0379)	(0.0404)	(0.0404)
Residual		0.2897	0.2697	0.2697
		(0.0370)	(0.0434)	(0.0437)
Residual × NonAgri		-0.4959	-0.4601	-0.4971
		(0.0414)	(0.0817)	(0.0826)
Residual^2			-0.0365	-0.0056
			(0.0482)	(0.0481)
Residual^2 × NonAgri			0.0188	0.0297
			(0.0948)	(0.0951)
Residual × Z				0.2422
				(0.0552)
Residual × NonAgri × Z				0.0784
				(0.0705)
First-stage specification		Linear + interactions with Z	Linear + interactions with Z	Linear + interactions with Z
Individual controls	Y	Y	Y	Y
Province × Year FE	Y	Y	Y	Y
Village FE	Y	Y	Y	Y
Observations	229,849	229,849	229,849	229,849
R-squared	–	0.4202	0.4202	0.4209
Kleibergen-Paap F-Stat	14.87	–	–	–

Notes: For the specification in Column (1), the dependent variables *Hukou Index: below median × NonAgri* and *Hukou Index: above median × NonAgri* are instrumented by *Hukou Index: below median × NRPS × Elder60* and *Hukou Index: above median × NRPS × Elder60*, respectively. The individual controls include age, age squared, years of education, a dummy for gender, a dummy for poor health, a dummy indicating whether there is an elderly member aged 60 or above in the household, and the share of months in year t that the NRPS has been in effect. The first-stage estimation of the selection equation for specifications in Columns (2)–(4) additionally includes the interactions between the IV ($NRPS \times Elder60$) and the individual characteristics (i.e., age, age squared, years of education, a dummy for gender, and a dummy for poor health). Robust standard errors are clustered at the village×year level.

TABLE C.2

NRPS AND SECTOR OF EMPLOYMENT:
BY GENDER AND THE PRESENCE OF CHILDREN AGED 15 OR BELOW

	(1)	(2)	(3)	(4)	(5)
Dep. Var.:	NonAgri Female	NonAgri Male	NonAgri All	NonAgri Female	NonAgri Male
Elder60 × NRPS × Child15			0.0153	0.0305	0.0010
			(0.0115)	(0.0153)	(0.0140)
Elder60 × NRPS	0.0692	0.0209	0.0313	0.0479	0.0197
	(0.0094)	(0.0086)	(0.0093)	(0.0125)	(0.0114)
Individual controls	Y	Y	Y	Y	Y
Province × Year FE	Y	Y	Y	Y	Y
Village FE	Y	Y	Y	Y	Y
Observations	108,041	121,807	229,849	108,041	121,807
R-squared	0.3625	0.3537	0.3630	0.3676	0.3544

Notes: Individual controls include age, age squared, years of education, a dummy for gender, and a dummy for poor health. Columns (3)–(5) also control for $Elder60 \times Child15$, $NRPS \times Child15$, $Elder60$, $NRPS$, and $Child15$ to account for their independent effects on sectoral choice. Robust standard errors are clustered at the village×year level.

TABLE C.3

NRPS AND SECTOR OF EMPLOYMENT:
BY EDUCATION GROUP AND AGE GROUP

	(1)	(2)	(3)	(4)	(5)
Dep. Var.:	NonAgri	NonAgri	NonAgri	NonAgri	NonAgri
Sample:	Years of Education			Age	
	$[0, 6]$	$(6, 9]$	$(9, 12]$	$[20, 39]$	$[40, 55]$
Elder60 × NRPS	0.0416	0.0349	0.0358	0.0245	0.0652
	(0.0111)	(0.0081)	(0.0174)	(0.0082)	(0.0103)
NRPS	-0.0029	0.0178	0.0227	0.0142	0.0068
	(0.0118)	(0.0112)	(0.0188)	(0.0133)	(0.0107)
Elder60	0.0119	0.0302	0.0239	0.0161	0.0233
	(0.0036)	(0.0033)	(0.0079)	(0.0034)	(0.0032)
Individual controls	Y	Y	Y	Y	Y
Province × Year FE	Y	Y	Y	Y	Y
Village FE	Y	Y	Y	Y	Y
Observations	85,636	125,438	18,766	116,540	113,308
R-squared	0.2614	0.3617	0.4421	0.3536	0.2112

Notes: Individual controls include age, age squared, years of education, a dummy for gender, and a dummy for poor health. Robust standard errors are clustered at the village×year level.

TABLE C.4

NRPS AND SECTOR OF EMPLOYMENT:
BY LOCATION AND AGE GROUP OF THE ELDERLY MEMBER

Dep. Var.:	(1) NonAgri within County	(2) NonAgri outside County within Province	(3) NonAgri outside Province	(4) NonAgri All
Elder60 × NRPS	0.0027	0.0150	0.0239	
	(0.0044)	(0.0049)	(0.0068)	
NRPS	0.0132	0.0043	-0.0064	0.0104
	(0.0069)	(0.0062)	(0.0068)	(0.0102)
Elder60	0.0024	0.0089	0.0120	
	(0.0015)	(0.0018)	(0.0022)	
Elder55–59×NRPS				0.0046
				(0.0080)
Elder60–69×NRPS				0.0512
				(0.0102)
Elder≥70×NRPS				0.0359
				(0.0088)
Elder55–59				0.0450
				(0.0032)
Elder60–69				0.0493
				(0.0037)
Elder≥70				0.0153
				(0.0031)
Individual controls	Y	Y	Y	Y
Province × Year FE	Y	Y	Y	Y
Village FE	Y	Y	Y	Y
Observations	229,849	229,849	229,849	229,849
R-squared	0.1473	0.1645	0.2892	0.3626

Notes: Individual controls include age, age squared, years of education, a dummy for gender, and a dummy for poor health. Robust standard errors are clustered at the village×year level.

TABLE C.5

NRPS AND ELDERLY LABOR SUPPLY OF THE ELDERLY

Dep. Var.: Working days	(1)	(2)	(3)	(4)	(5)
Sample:	All	Age<70	Age≥70	Good Health	Poor Health
	Poisson	Poisson	Poisson	Poisson	Poisson
NRPS	-0.0837	-0.0932	0.1168	-0.1283	0.0750
	(0.0384)	(0.0383)	(0.0830)	(0.0439)	(0.0674)
Individual controls	Y	Y	Y	Y	Y
Province × Year FE	Y	Y	Y	Y	Y
Village FE	Y	Y	Y	Y	Y
Observations	46,984	29,308	17,068	23,680	23,094

Notes: Individual controls include years of education, a dummy for gender, and dummies of health status. Robust standard errors are clustered at the village×year level.

TABLE C.6

NRPS AND OTHER HOUSEHOLD-LEVEL OUTCOMES

	(1)	(2)	(3)	(4)	(5)
Dep. Var.:	Savings Rate	ln(1+Fixed Investment)	ln(1+Loan)	ΔArable Land	Immediate Input per Labor Input
	OLS	OLS	OLS	OLS	OLS
Elder60×NRPS	0.0052	-0.0215	0.0297	-0.0009	0.0179
	(0.0080)	(0.0359)	(0.0452)	(0.0042)	(0.0208)
NRPS	0.0028	-0.0468	-0.0704	-0.0151	-0.0280
	(0.0134)	(0.0648)	(0.0866)	(0.0079)	(0.0392)
Elder60	-0.0035	0.0100	-0.1294	-0.0078	-0.0060
	(0.0035)	(0.0153)	(0.0201)	(0.0016)	(0.0068)
Individual controls	Y	Y	Y	Y	Y
Province×Year FE	Y	Y	Y	Y	Y
Village FE	Y	Y	Y	Y	Y
Observations	107,911	110,743	110,743	103,076	101,474
R-squared	0.1201	0.0890	0.1320	0.0848	0.4475

Notes: Household controls include average age, average years of education, the share of males, the share of working-age household members who have poor health, and start-of-period arable land per capita. Robust standard errors are clustered at the village×year level.

TABLE C.7

NRPS, MIGRATION, AND HOUSEHOLD WEALTH

Dep. Var.: NonAgri Wealth Measure:	(1) Deposits	(2) Deposits	(3) Cash	(4) Cash
Wealth: Below Median×Elder60×NRPS (β_1)	0.0586		0.0431	
	(0.0143)		(0.0140)	
Wealth: Above Median×Elder60×NRPS (β_2)	0.0581		0.0640	
	(0.0124)		(0.0110)	
Bottom Tercile×Elder60×NRPS (π_1)		0.0572		0.0513
		(0.0203)		(0.0167)
Middle Tercile×Elder60×NRPS (π_2)		0.0633		0.0402
		(0.0140)		(0.0142)
Top Tercile×Elder60×NRPS (π_3)		0.0518		0.0690
		(0.0147)		(0.0128)
Individual and household controls	Y	Y	Y	Y
Province×Year FE	Y	Y	Y	Y
Village FE	Y	Y	Y	Y
Observations	117,297	117,297	172,446	172,446
R-squared	0.3859	0.3863	0.3600	0.3603
F test	$\beta_1 = \beta_2$	$\pi_1 = \pi_2 = \pi_3$	$\beta_1 = \beta_2$	$\pi_1 = \pi_2 = \pi_3$
p-value	0.979	0.840	0.214	0.270

Notes: All regressions control for wealth group dummies, the interaction terms of wealth group dummies and Elder60, and the interaction terms of wealth group dummies and NRPS. Individual controls include age, age squared, years of education, a dummy for gender, and a dummy for poor health. Robust standard errors are clustered at the village×year level.

TABLE C.8

SECTOR OF EMPLOYMENT AND DAILY WAGE
(HUMAN-CAPITAL-ADJUSTED MEASURE)

Dep. Var.: ln Daily Wage	(1) OLS	(2) IV	(3) CF
Migration	0.2821	0.8781	0.2887
	(0.0140)	(0.3535)	(0.0282)
Individual controls	Y	Y	Y
Province× Year FE	Y	Y	Y
Village FE	Y	Y	Y
Observations	229,849	229,849	229,236
R-squared	0.4247	–	0.4257
Kleibergen-Paap F-Stat	–	31.02	–

Notes: Columns (1)–(3) respectively re-estimate the specifications from columns (1), (3), and (5) of Table 3. Individual controls include age, age squared, years of education, a dummy for gender, a dummy for poor health, a dummy indicating whether there is an elderly aged 60 or above residing in the household, and the share of months in year t that the NRPS has been in effect. Robust standard errors are clustered at the village×year level.

APPENDIX D: A COLLECTIVE DECISION MODEL

Now, we consider the optimal labor allocation when family members can pool their incomes and act cooperatively. In this case, for $j = a$, the allocation problem is:

$$V_{c,a} = \max_{l_{o,a}, l_y} \left\{ \ln\left(\frac{w_a(h_o l_{o,a} + h_a l_y) + p_a T}{2}\right) - \frac{1}{1+\frac{1}{\phi}} \left(\xi l_{o,a} + l_y\right)^{1+\frac{1}{\phi}} \right\}$$

For $j = na$, the allocation problem is

$$V_{c,na} = \max_{l_{o,na}, l_{na}} \left\{ \ln\left(\frac{w_a h_o l_{o,na} + w_{na} h_{na}(l_{na} - m_0) + p_a T}{2}\right) - \frac{1}{1+\frac{1}{\phi}} \left(\xi l_{o,na} + l_{na}\right)^{1+\frac{1}{\phi}} \right\}$$

Comparing the collective model with the non-cooperative Nash game, we formulate three propositions:

Proposition 3 *In the collective model, the adult children's labor supply declines with the size of the pension transfer T.*

Proposition 4 *When elderly parents and adult children play a non-cooperative Nash game, there exists a transfer level T where $\tilde{T} = \xi^{-1}$, such that the household will achieve the optimal labor allocation in the collective model without transfers.*

Proposition 5 *For any level of transfer T, the collective model always yields higher joint household welfare compared with the non-cooperative Nash game.*

The proofs of these three propositions are in the next Section.

APPENDIX E: THEORY REPLICATION APPENDIX

E.1. *Technical Details of the General Equilibrium Model*

E.1.1. *Household Production Income*

This section derives the formula for household production income. Given the total labor supply of old and young agents in the rural area, l_o and l_y, and the output prices p_a and p_{na}, the household allocates labor between agriculture and non-agriculture to maximize total household income:

$$\max_{l_{oa,t},l_{or,t},l_{ya,t},l_{yr,t}} \left\{ p_{a,t}A_{a,t}\left(h_{o,t}l_{oa,t}+h_{y,t}l_{ya,t}\right)^{\alpha}+p_{na,t}A_{r,t}\left(h_{o,t}l_{or,t}+h_{y,t}l_{yr,t}\right)^{\alpha} \right\}$$

subject to

$$l_{ij,t}\geq 0, i=o,y, j=a,r;$$

$$l_{ia,t}+l_{ir,t}=l_{i,t}, i=o,y.$$

The F.O.Cs are

$$\alpha p_{a,t}A_{a,t}\left(h_{o,t}l_{oa,t}+h_{y,t}l_{ya,t}\right)^{\alpha-1}h_{o,t}=\alpha p_{na,t}A_{r,t}\left(h_{o,t}l_{or,t}+h_{y,t}l_{yr,t}\right)^{\alpha-1}h_{o,t}=\lambda_{o,t}$$

$$\alpha p_{a,t}A_{a,t}\left(h_{o,t}l_{oa,t}+h_{y,t}l_{ya,t}\right)^{\alpha-1}h_{y,t}=\alpha p_{na,t}A_{r,t}\left(h_{o,t}l_{or,t}+h_{y,t}l_{yr,t}\right)^{\alpha-1}h_{y,t}=\lambda_{y,t}$$

Thus, we have

$$h_{o,t}l_{oa,t}+h_{y,t}l_{ya,t}=\left(\frac{p_{a,t}A_{a,t}}{p_{na,t}A_{r,t}}\right)^{\frac{1}{1-\alpha}}\left(h_{o,t}l_{or,t}+h_{y,t}l_{yr,t}\right)$$

From the budget constraints, we have

$$h_{o,t}l_{o,t} + h_{y,t}l_{y,t} = \left[1 + \left(\frac{p_{a,t}A_{a,t}}{p_{na,t}A_{r,t}}\right)^{\frac{1}{1-\alpha}}\right](h_{o,t}l_{or,t} + h_{y,t}l_{yr,t}),$$

which implies

$$h_{o,t}l_{or,t} + h_{y,t}l_{yr,t} = \frac{(p_{na,t}A_{r,t})^{\frac{1}{1-\alpha}}}{(p_{a,t}A_{a,t})^{\frac{1}{1-\alpha}} + (p_{na,t}A_{r,t})^{\frac{1}{1-\alpha}}}(h_{o,t}l_{o,t} + h_{y,t}l_{y,t}),$$

$$h_{o,t}l_{oa,t} + h_{y,t}l_{ya,t} = \frac{(p_{a,t}A_{a,t})^{\frac{1}{1-\alpha}}}{(p_{a,t}A_{a,t})^{\frac{1}{1-\alpha}} + (p_{na,t}A_{r,t})^{\frac{1}{1-\alpha}}}(h_{o,t}l_{o,t} + h_{y,t}l_{y,t}).$$

Therefore, the agricultural and non-agricultural output of the household are

$$y_{a,t} = A_{a,t}(h_{o,t}l_{oa,t} + h_{y,t}l_{ya,t})^{\alpha} = \frac{A_{a,t}(p_{a,t}A_{a,t})^{\frac{\alpha}{1-\alpha}}}{\left[(p_{a,t}A_{a,t})^{\frac{1}{1-\alpha}} + (p_{na,t}A_{r,t})^{\frac{1}{1-\alpha}}\right]^{\alpha}}(h_{o,t}l_{o,t} + h_{y,t}l_{y,t})^{\alpha}$$

$$y_{r,t} = A_{r,t}(h_{o,t}l_{or,t} + h_{y,t}l_{yr,t})^{\alpha} = \frac{A_{r,t}(p_{na,t}A_{r,t})^{\frac{\alpha}{1-\alpha}}}{\left[(p_{a,t}A_{a,t})^{\frac{1}{1-\alpha}} + (p_{na,t}A_{r,t})^{\frac{1}{1-\alpha}}\right]^{\alpha}}(h_{o,t}l_{o,t} + h_{y,t}l_{y,t})^{\alpha}$$

And the household production income is

$$y_{f,t} = \left[(p_{a,t}A_{a,t})^{\frac{1}{1-\alpha}} + (p_{na,t}A_{r,t})^{\frac{1}{1-\alpha}}\right]^{1-\alpha}(h_{o,t}l_{o,t} + h_{y,t}l_{y,t})^{\alpha} = A_{f,t}h_{f,t}^{\alpha},$$

where

$$A_{f,t} = \left[(p_{a,t}A_{a,t})^{\frac{1}{1-\alpha}} + (p_{na,t}A_{r,t})^{\frac{1}{1-\alpha}}\right]^{1-\alpha}, \quad \text{and} \quad h_{f,t} = h_{o,t}l_{o,t} + h_{y,t}l_{y,t}.$$

E.1.2. *Labor Supply Decisions of Rural Households*

The Case of No Migration This section derives the first-order conditions for parents' and adult children's optimization problems when adult children choose not to migrate.

The parent's optimization problem is

$$\max_{l_o \in [0, n_o]} n_o \left\{ \frac{1}{1-\gamma} c\left(e_o\right)^{1-\gamma} - \frac{\eta}{1+\dfrac{1}{\phi}} \frac{\left(\xi l_o + l_y\right)^{1+\frac{1}{\phi}}}{\left(n_o + n_y\right)^{1+\frac{1}{\phi}}} \right\}.$$

Note that

$$\frac{h_o l_o}{h_f} y_f = h_o l_o A_f \left(h_o l_o + h_y l_y\right)^{\alpha - 1}.$$

So,

$$\begin{aligned} \tfrac{\partial}{\partial l_o}\left(\tfrac{h_o l_o}{h_f} y_f\right) &= h_o A_f \left[\left(h_o l_o + h_y l_y\right)^{\alpha-1} - (1-\alpha) h_o l_o \left(h_o l_o + h_y l_y\right)^{\alpha-2}\right] \\ &= h_o A_f \left(h_o l_o + h_y l_y\right)^{\alpha-2} \left(\alpha h_o l_o + h_y l_y\right). \end{aligned}$$

Thus, the F.O.C. for l_o is

$$c^{-\gamma} c'\left(e_o\right) \frac{h_o A_f}{n_o} \left(h_o l_o + h_y l_y\right)^{\alpha-2} \left(\alpha h_o l_o + h_y l_y\right) / \kappa_r = \eta \xi \frac{\left(\xi l_o + l_y\right)^{\frac{1}{\phi}}}{\left(n_o + n_y\right)^{1+\frac{1}{\phi}}}. \tag{E.1}$$

Similarly, the child's optimization problem is

$$V_a = \max_{l_y \in [0, n_y]} n_y \left\{ \frac{1}{1-\gamma} c\left(e_y\right)^{1-\gamma} - \frac{\eta}{1+\dfrac{1}{\phi}} \frac{\left(\xi l_o + l_y\right)^{1+\frac{1}{\phi}}}{\left(n_o + n_y\right)^{1+\frac{1}{\phi}}} \right\},$$

and the corresponding F.O.C. for l_y is

$$c^{-\gamma}c'\left(e_y\right)\frac{h_yA_f}{n_y}\left(h_ol_o+h_yl_y\right)^{\alpha-2}\left(\alpha h_yl_y+h_ol_o\right)/\kappa_r=\eta\frac{\left(\xi l_o+l_y\right)^{\frac{1}{\phi}}}{\left(n_o+n_y\right)^{1+\frac{1}{\phi}}}. \tag{E.2}$$

The equations (E.1) and (E.2) can be used to jointly solve for l_o and l_y when the child does not migrate.

The Case of Migration In this section, we derive the first-order conditions for parents' and adult children's optimization problems when adult children choose to migrate.

The parent's optimization problem in this case is

$$\max_{l_o\in[0,n_o]} n_o\left\{\frac{1}{1-\gamma}c\left(e_o\right)^{1-\gamma}-\frac{\eta}{1+\dfrac{1}{\phi}}\frac{\left(\xi l_o+l_y+l_{na}\right)^{1+\frac{1}{\phi}}}{\left(n_o+n_y\right)^{1+\frac{1}{\phi}}}\right\},$$

and the F.O.C. for l_o is

$$c^{-\gamma}c'\left(e_o\right)\frac{h_o}{n_o}A_f\left(h_ol_o+h_yl_y\right)^{\alpha-2}\left(\alpha h_ol_o+h_yl_y\right)/\kappa_r=\eta\xi\frac{\left(\xi l_o+l_y+l_{na}\right)^{\frac{1}{\phi}}}{\left(n_o+n_y\right)^{1+\frac{1}{\phi}}} \tag{E.3}$$

The child's optimization problem is

$$V_{na}=\max_{l_y,l_{na}\in[0,n_y],l_y+l_{na}\leq n_y} n_y\left\{\frac{1}{1-\gamma}c\left(e_y\right)^{1-\gamma}-\frac{\eta}{1+\dfrac{1}{\phi}}\frac{\left(\xi l_o+l_y+l_{na}\right)^{1+\frac{1}{\phi}}}{\left(n_o+n_y\right)^{1+\frac{1}{\phi}}}\right\},$$

and the corresponding F.O.C.s for l_y and l_{na} are

$$c^{-\gamma}c'(e_y)\frac{h_yA_f}{n_y}\left(h_ol_o+h_yl_y\right)^{\alpha-2}\left(\alpha h_yl_y+h_ol_o\right)/\kappa_r=\eta\frac{\left(\xi l_o+l_y+l_{na}\right)^{\frac{1}{\phi}}}{\left(n_o+n_y\right)^{1+\frac{1}{\phi}}} \tag{E.4}$$

and

$$c^{-\gamma}c'(e_y)\frac{w_{na}h_{na}}{n_y}(1-m_1)/\kappa_u = \eta\frac{(\xi l_o + l_y + l_{na})^{\frac{1}{\phi}}}{(n_o+n_y)^{1+\frac{1}{\phi}}} \tag{E.5}$$

The Equations (E.3), (E.4), and (E.11) can be used to jointly solve for l_o, l_y, and l_{na} when the child migrates. Comparing (E.4) and (E.11), if $\max_{l_y\in[0,n_y]}\{h_yA_f\left(h_ol_o+h_yl_y\right)^{\alpha-2}\left(\alpha h_yl_y+h_ol_o\right)/\kappa_r\} < w_{na}h_{na}(1-m_1)/\kappa_u$, then $l_y=0$.

E.2. *Technical Details of the Illustrative Nash Equilibrium Model*

This subsection details the proofs of the propositions discussed in Section 4.3 of the main text.

E.2.1. *Proof of Proposition 1*

The parent's labor supply decision problem is

$$V_{o,j} = \max_{l_{o,j}\in[0,1]}\left\{\ln\left(w_ah_ol_{o,j}+p_aT\right) - \frac{1}{1+\dfrac{1}{\phi}}\left(\xi l_{o,j} + l_y\mathbf{1}_{\{j=a\}} + l_{na}\mathbf{1}_{\{j=na\}}\right)^{1+\frac{1}{\phi}}\right\}.$$

The interior solution to the problem should satisfy the following F.O.C.:

$$\frac{1}{l_{o,j}+\tilde{T}} = \xi\left(\xi l_{o,j} + l_y\mathbf{1}_{\{j=a\}} + l_{na}\mathbf{1}_{\{j=na\}}\right)^{\frac{1}{\phi}}. \tag{E.6}$$

where $\tilde{T}=\frac{p_aT}{w_ah_o}$.

Case of No Migration For $j=a$, the children's labor supply problem is

$$V_a = \max_{l_y\in[0,1]}\left\{\ln\left(w_ah_al_y\right) - \frac{1}{1+\dfrac{1}{\phi}}\left(\xi l_{o,a}+l_y\right)^{1+\frac{1}{\phi}}\right\},$$

which has the following F.O.C.:

$$\frac{1}{l_y} = (\xi l_{o,a} + l_y)^{\frac{1}{\phi}} . \tag{E.7}$$

Comparing (E.7) with (E.6) for $j = a$, we have

$$l_y = \xi \left(l_{o,a} + \tilde{T} \right) . \tag{E.8}$$

Substituting (E.8) into (E.7) yields the following:

$$1 = l_y \left(2l_y - \xi \tilde{T} \right)^{\frac{1}{\phi}} . \tag{E.9}$$

In addition, substituting (E.8) into (E.6) yields the following

$$1 = \xi^{1+\frac{1}{\phi}} \left(l_{o,a} + \tilde{T} \right) \left(2l_{o,a} + \tilde{T} \right)^{\frac{1}{\phi}} . \tag{E.10}$$

Let $l^*_{o,a}(\tilde{T})$ and $l^*_y(\tilde{T})$ be the solutions to (E.9) and (E.10). It can be easily shown that

$$\frac{\partial l^*_y}{\partial \tilde{T}} > 0, \quad \text{and} \quad \frac{\partial l^*_{o,a}}{\partial \tilde{T}} < 0.$$

We now consider the conditions when the optimal labor allocation is indeed interior. For $l^*_{o,a}(\tilde{T}) > 0$, from (E.10) we can see that this requires the following to be true: $1 > \xi^{1+\frac{1}{\phi}} \tilde{T}^{1+\frac{1}{\phi}}$ or $\tilde{T} < \xi^{-1}$. Under this condition, we can see from (E.9) that $l^*_y(\tilde{T})$ also satisfies the interior condition. For $\tilde{T} \geq \xi^{-1}$, we have $l_{o,a} = 0$. In this case, the children's problem becomes

$$V_a = \max_{l_y \in [0,1]} \left\{ \ln (w_a h_a l_y) - \frac{1}{1 + \dfrac{1}{\phi}} l_y^{1+\frac{1}{\phi}} \right\},$$

which implies $l_y = 1$. That is, for $\tilde{T} \geq \xi^{-1}$, there is complete specialization; $l_{o,a} = 0$ and $l_y = 1$.

To summarize, in the case of no-migration, the household's labor allocation $(l_y, l_{o,a})$ is given by the solutions to (E.9) and (E.10) if $\tilde{T} < \xi^{-1}$. If $\tilde{T} \geq \xi$, then, $l_y = 1$ and $l_{o,a} = 0$.

Case of Migration For $j = na$, the children's labor supply decision problem is

$$V_{na} = \max_{l_{na} \in [0,1]} \left\{ \ln\left(w_{na} h_{na}(l_{na} - m_0)\right) - \frac{1}{1 + \frac{1}{\phi}} \left(\xi l_{o,na} + l_{na}\right)^{1+\frac{1}{\phi}} \right\},$$

which has the following F.O.C.:

$$\frac{1}{l_{na} - m_0} = \left(\xi l_{o,na} + l_{na}\right)^{\frac{1}{\phi}}. \tag{E.11}$$

Comparing (E.11) with (E.6) for $j = na$, we have

$$l_{na} - m_0 = \xi \left(l_{o,na} + \tilde{T} \right). \tag{E.12}$$

Substituting (E.12) into (E.11) yields the following:

$$1 = (l_{na} - m_0) \left(2 l_{na} - m_0 - \xi \tilde{T} \right)^{\frac{1}{\phi}}. \tag{E.13}$$

In addition, substituting (E.12) into (E.6) yields the following

$$1 = \xi^{1+\frac{1}{\phi}} \left(l_{o,na} + \tilde{T} \right) \left(2 l_{o,na} + \tilde{T} + \frac{m_0}{\xi} \right)^{\frac{1}{\phi}}. \tag{E.14}$$

Let $l^*_{na}(m_0, \tilde{T})$ and $l^*_{o,na}(m_0, \tilde{T})$ be the solutions to (E.13) and (E.14). It is clear that

$$\frac{\partial l^*_{na}}{\partial \tilde{T}} > 0, \quad \text{and} \quad \frac{\partial l^*_{o,na}}{\partial \tilde{T}} < 0.$$

From (E.13), we can also show that

$$\frac{\partial l_{na}^*}{\partial \tilde{m_0}} > 0, .$$

Comparing (E.7) and (E.13), we have

$$l_{na}^*(0,\tilde{T}) = l_y^*(\tilde{T}).$$

Again, consider the interior solution condition $l_{o,na} > 0$, from (E.14) we need $1 > \xi\tilde{T}\left(\xi\tilde{T} + m_0\right)^{\frac{1}{\phi}}$. Let $\tilde{T}^*(m_0)$ be the value of $\tilde{T}$ such that

$$\xi\tilde{T}\left(\xi\tilde{T} + m_0\right)^{\frac{1}{\phi}} = 1. \tag{E.15}$$

Clearly $\tilde{T}^*(m_0) \leq \xi^{-1}$ and the equality holds if and only if $m_0 = 0$. Thus, a necessary condition for $l_{o,na} > 0$ is $\tilde{T} < \tilde{T}^*(m_0)$. In this case, if $l_{na}^*(m_0,\tilde{T}) < 1$, then $(l_{na}^*(m_0,\tilde{T}), l_{o,na}^*(m_0,\tilde{T}))$ is the household's labor allocation.

If $\tilde{T} < \tilde{T}^*(m_0)$ but $l_{na}^*(m_0,\tilde{T}) \geq 1$, then we have the corner solution for children's labor allocation, $l_{na} = 1$. In this case, the parents' problem becomes

$$V_{o,na} = \max_{l_{o,na}\in[0,1]} \left\{ \ln\left(w_a h_o l_{o,na} + p_a T\right) - \frac{1}{1+\dfrac{1}{\phi}}\left(\xi l_{o,j} + 1\right)^{1+\frac{1}{\phi}} \right\},$$

which has the following F.O.C.

$$1 = \xi\left(l_{o,na} + \tilde{T}\right)\left(\xi l_{o,na} + 1\right)^{\frac{1}{\phi}}. \tag{E.16}$$

For (E.16) to have an interior solution, we need $\tilde{T}\xi < 1$, which holds when $\tilde{T} < \tilde{T}^*(m_0)$.

For $\tilde{T} > \tilde{T}^*(m_0)$, we have $l_{o,na} = 0$. In this case, the children's problem becomes

$$V_{na} = \max_{l_{na}\in[0,1]} \left\{ \ln\left(w_{na}h_{na}(l_{na} - m_0)\right) - \frac{1}{1+\frac{1}{\phi}} l_{na}^{1+\frac{1}{\phi}} \right\},$$

Note that

$$\partial V_{na}/\partial l_{na} = \frac{1}{l_{na} - m_0} - l_{na}^{\frac{1}{\phi}} \geq \frac{1}{1-m_0} - 1 > 0$$

for any $m_0 > 0$. Thus, the optimal l_{na} for the children is $l_{na} = 1$.

To summarize, if $\tilde{T} < \tilde{T}^*(m_0)$, then, $l_{na} = \min\left\{1, l_{na}^*(m_0, \tilde{T})\right\}$, and $l_{0,na}$ is the solution to the following equation:

$$1 = \xi\left(l_{o,na} + \tilde{T}\right)\left(\xi l_{o,na} + \min\left\{1, l_{na}^*(m_0, \tilde{T})\right\}\right)^{\frac{1}{\phi}}. \qquad \text{(E.17)}$$

If $\tilde{T} \geq \tilde{T}^*(m_0)$, then $l_{na} = 1$ and $l_{o,na} = 0$.

E.2.2. *Proof of Proposition 2*

In general,

$$M(m_0, \tilde{T}) = \ln\left(\frac{l_y}{l_{na} - m_0}\right) - \frac{1}{1+\frac{1}{\phi}}\left[(\xi l_{o,a} + l_y)^{1+\frac{1}{\phi}} - (\xi l_{o,na} + l_{na})^{1+\frac{1}{\phi}}\right]. \qquad \text{(E.18)}$$

Proof of $M(0, \tilde{T}) = 0$. When $m_0 = 0$, from (E.6), (E.7), and (E.11), we have $l_{o,a} = l_{o,na}$ and $l_y = l_{na}$. Therefore, according to equation (E.18), $M(m_0, \tilde{T}) = 0$.

Proof of $\partial M(m_0,\tilde{T})/\partial m_0 > 0$. A change in m_0 will affect l_{na} but not l_y. Define $\tilde{l}_{na} = l_{na} - m_0$, using (E.7) and (E.11), equation (E.18) can be re-written as

$$M(m_0,\tilde{T}) = \ln\left(\frac{l_y}{\tilde{l}_{na}}\right) - \frac{1}{1+\frac{1}{\phi}}\left[\left(\frac{1}{l_y}\right)^{1+\phi} - \left(\frac{1}{\tilde{l}_{na}}\right)^{1+\phi}\right].$$

Therefore, taking the derivative with respect to m_0, we have

$$\frac{\partial M(m_0,\tilde{T})}{\partial m_0} = -\left(\frac{1}{\tilde{l}_{na}} + \phi\left(\frac{1}{\tilde{l}_{na}}\right)^{\phi}\left(\tilde{l}_{na}\right)^{-2}\right)\frac{\partial \tilde{l}_{na}}{\partial m_0}$$

From (E.14), we have $\frac{\partial l_{o,na}}{\partial m_0} < 0$. Then from (E.12), we have $\tilde{l}_{na} = \xi\left(l_{o,na} + \tilde{T}\right)$. Therefore, $\frac{\partial \tilde{l}_{na}}{\partial m_0} < 0$, and $\frac{\partial M(m_0,\tilde{T})}{\partial m_0} > 0$.

Proof of $\partial M(m_0,\tilde{T})/\partial \tilde{T} < 0$. From the envelope theorem, we have

$$\frac{\partial(V_{na} - V_a)}{\partial T} = -\frac{\partial M(m_0,\tilde{T})}{\partial \tilde{T}} = -\xi\left(\xi l_{o,na} + l_{na}\right)^{\frac{1}{\phi}}\frac{\partial l_{o,na}}{\partial \tilde{T}} + \xi\left(\xi l_{o,a} + l_y\right)^{\frac{1}{\phi}}\frac{\partial l_{o,a}}{\partial \tilde{T}}.$$

From (E.10) and (E.14), we have

$$-\frac{\partial M(m_0,\tilde{T})}{\partial \tilde{T}} = -\frac{\frac{\partial l_{o,na}}{\partial \tilde{T}}}{l_{o,na} + \tilde{T}} + \frac{\frac{\partial l_{o,a}}{\partial \tilde{T}}}{l_{o,a} + \tilde{T}}. \tag{E.19}$$

Take $\log$ and differentiate (E.10) and (E.14) with respect to $\tilde{T}$, we have

$$\frac{\frac{\partial l_{o,a}}{\partial \tilde{T}} + 1}{l_{o,a} + \tilde{T}} + \frac{2\frac{\partial l_{o,a}}{\partial \tilde{T}} + 1}{\phi\left(2l_{o,a} + \tilde{T}\right)} = 0;$$

$$\frac{\frac{\partial l_{o,na}}{\partial \tilde{T}}+1}{l_{o,na}+\tilde{T}}+\frac{2\frac{\partial l_{o,a}}{\partial \tilde{T}}+1}{\phi\left(2l_{o,a}+\tilde{T}+\frac{m_0}{\xi}\right)}=0.$$

Therefore,

$$\frac{\partial l_{o,a}}{\partial \tilde{T}}=-\frac{(2\phi+1)l_{o,a}+(1+\phi)\tilde{T}}{2(1+\phi)l_{o,a}+(2+\phi)\tilde{T}}, \tag{E.20}$$

$$\frac{\partial l_{o,na}}{\partial \tilde{T}}=-\frac{(2\phi+1)l_{o,na}+(1+\phi)\tilde{T}+\phi\frac{m_0}{\xi}}{2(1+\phi)l_{o,na}+(2+\phi)\tilde{T}+\phi\frac{m_0}{\xi}}, \tag{E.21}$$

It can be easily shown that, for $m>0$,

$$\frac{(2\phi+1)l_{o,na}+(1+\phi)\tilde{T}+\phi\frac{m_0}{\xi}}{2(1+\phi)l_{o,na}+(2+\phi)\tilde{T}+\phi\frac{m_0}{\xi}}>\frac{(2\phi+1)l_{o,na}+(1+\phi)\tilde{T}}{2(1+\phi)l_{o,na}+(2+\phi)\tilde{T}}.$$

Therefore, from (E.21), we have

$$\frac{\frac{\partial l_{o,na}}{\partial \tilde{T}}}{l_{o,na}+\tilde{T}}<-\frac{(2\phi+1)l_{o,na}+(1+\phi)\tilde{T}}{\left(l_{o,na}+\tilde{T}\right)\left(2(1+\phi)l_{o,na}+(2+\phi)\tilde{T}\right)}$$

Let

$$f(x)=\frac{(2\phi+1)x+(1+\phi)\tilde{T}}{\left(x+\tilde{T}\right)\left(2(1+\phi)x+(2+\phi)\tilde{T}\right)}$$

If we can show that $f'(x)<0$, then, since $l_{o,na}<l_{o,a}$, we have $f(l_{o,na})>f(l_{o,a})$. Thus,

$$\frac{\frac{\partial l_{o,na}}{\partial \tilde{T}}}{l_{o,na}+\tilde{T}}<-f(l_{o,na})<-f(l_{0,a})=\frac{\frac{\partial l_{o,a}}{\partial \tilde{T}}}{l_{o,a}+\tilde{T}},$$

and $\partial M(m_0,\tilde{T})/\partial \tilde{T}<0$.

To prove that $f(x)$ is a decreasing function is equivalent to proving that $\ln f(x)$ is a decreasing function.

$$\begin{aligned}\frac{\partial \ln f(x)}{\partial x} &= \frac{2\phi+1}{(2\phi+1)x+(1+\phi)\tilde{T}} - \frac{1}{x+\tilde{T}} - \frac{2(1+\phi)}{2(1+\phi)x+(2+\phi)\tilde{T}} \\ &= \frac{\phi\tilde{T}}{\left(x+\tilde{T}\right)\left[(2\phi+1)x+(1+\phi)\tilde{T}\right]} - \frac{2(1+\phi)}{2(1+\phi)x+(2+\phi)\tilde{T}} \\ &< \frac{\phi}{(2\phi+1)x+(1+\phi)\tilde{T}} - \frac{2(1+\phi)}{2(1+\phi)x+(2+\phi)\tilde{T}} \\ &= \frac{-2(1+\phi)^2x-(2+2\phi+\phi^2)\tilde{T}}{\left[(2\phi+1)x+(1+\phi)\tilde{T}\right]\left[2(1+\phi)x+(2+\phi)\tilde{T}\right]} < 0.\end{aligned}$$

E.3. *Technical Details of the Collective Decision Model*

This subsection details the proofs of the propositions discussed in Appendix C.

E.3.1. *Proof of Proposition 3*

We assume that $\xi > \frac{h_o}{h_a}$, i.e., the parent has a comparative advantage in home production. Then, for $j = a$, it is optimal to set $l_{o,a} = 0$, and the optimal l_y satisfies the following F.O.C.

$$\frac{w_a h_a}{w_a h_a l_y + p_a T} = l_y^{\frac{1}{\phi}},$$

which implies that $dl_y/dT < 0$.

For $j = na$, again the optimal choice of parents' labor supply is $l_{o,na} = 0$. The optimal l_{na} satisfies the following F.O.C.

$$\frac{w_a h_a}{w_a h_a (l_{na} - m_o) + p_a T} = l_{na}^{\frac{1}{\phi}},$$

which implies that $dl_{na}/dT < 0$.

E.3.2. *Proof of Proposition 4*

Now we consider a case when $T = 0$ in the collective model and solve its optimal labor allocation. We assume that $\xi > \frac{h_o}{h_a}$, i.e., the parent has a comparative advantage in home production. Then, it is optimal to set $l_{o,a} = 0$, and the optimal l_y satisfies the following F.O.C.

$$\frac{1}{l_y} = l_y^{\frac{1}{\phi}},$$

which implies that the optimal l_y is 1. That is, the optimal household allocation is complete specialization. From our discussion in Section 4.3 of the main text, the household will reach the same decision in the non-cooperative case if the government chooses the transfer such that $\tilde{T} \geq \xi^{-1}$.

For $j = na$, again the optimal choice of parents' labor supply is $l_{o,na} = 0$. Thus,

$$V_{c,na} = \max_{l_{na}} \left\{ \ln \left(\frac{w_{na} h_{na}(l_{na} - m_0)}{2} \right) - \frac{1}{1 + \frac{1}{\phi}} l_{na}^{1+\frac{1}{\phi}} \right\}$$

Note that

$$\partial V_{c,na} / \partial l_{na} = \frac{1}{l_{na} - m_0} - l_{na}^{\frac{1}{\phi}} > \frac{1}{1 - m_0} - 1 > 0$$

for any $l_{na} < 1$. Therefore, choosing $l_{na} = 1$ is optimal. That is, the household's collective labor supply decision is complete specialization in this case as well. As we have shown above, as long as $\tilde{T} \geq \tilde{T}^*(m_0)$, the household's non-cooperative labor supply decision is also complete specialization.

Based on equation (E.15), $\tilde{T}^*(m_0) < \xi^{-1}$. Therefore, if the government chooses T such that $\tilde{T} = \xi^{-1}$, the household will achieve the optimal labor allocation even when parents and children play a non-cooperative Nash game. In this case, no matter whether children choose migration or not, we always have complete specialization within the household.

E.3.3. *Proof of Proposition 5*

When $j = a$, let $\bar{l}_{o,a}$ and $\bar{l}_y$ be the optimal labor supply of parents and children in the non-cooperative Nash equilibrium. The concavity of the utility function yields the following inequality:

$$\frac{1}{2}\left(\ln(w_a h_o \bar{l}_{o,a} + p_a T) + \ln(w_a h_a \bar{l}_y)\right) < \ln\left(\frac{w_a(h_o \bar{l}_{o,a} + h_a \bar{l}_y) + p_a T}{2}\right)$$

Therefore, we have

$$\begin{aligned} V_{o,a} + V_a &= \ln\left(w_a h_o \bar{l}_{o,j} + p_a T\right) - \frac{1}{1+\frac{1}{\phi}}\left(\xi \bar{l}_{o,j} + \bar{l}_y\right)^{1+\frac{1}{\phi}} + \ln\left(w_a h_a \bar{l}_y\right) - \frac{1}{1+\frac{1}{\phi}}\left(\xi \bar{l}_{o,a} + \bar{l}_y\right)^{1+\frac{1}{\phi}} \\ &\leq 2\left[\ln\left(\frac{w_a(h_o \bar{l}_{o,a} + h_a \bar{l}_y) + p_a T}{2}\right) - \frac{1}{1+\frac{1}{\phi}}\left(\xi \bar{l}_{o,a} + \bar{l}_y\right)^{1+\frac{1}{\phi}}\right] \\ &\leq 2\max_{l_{o,a}, l_y}\left\{\ln\left(\frac{w_a(h_o l_{o,a} + h_a l_y) + p_a T}{2}\right) - \frac{1}{1+\frac{1}{\phi}}\left(\xi l_{o,a} + l_y\right)^{1+\frac{1}{\phi}}\right\} \\ &= 2V_{c,a} \end{aligned}$$

Similarly, we can prove that when $j = na$, we have

$$V_{o,na} + V_{na} \leq 2V_{c,na}$$

Therefore, the optimal labor allocation in the collective model always yields higher joint household utility compared with that in the non-cooperative Nash game.

APPENDIX F: CALIBRATION APPENDIX

F.1. *Sector Prices and Distributional Costs*

This subsection outlines the calibration procedure for $p_{a,t}$ and $p_{na,t}$. Firstly, we construct prices for the benchmark year 2005 using the data from the GGDC productivity level database. Specifically, the GGDC provides sector-specific price data, denoted as $p_{i,2005}^{GGDC}$, for 10 sectors. These prices are defined as follows:

$$p_{i,2005}^{GGDC} = \frac{\text{National Price Value-added of Sector i and Year 2005}/E_{2005}}{\text{2005 International Price Value-added of Sector i and Year 2005}}, i = 1, ..., 10.$$

Here, E_{2005} is the China-US nominal exchange rate (the price of US dollar in RMB) in 2005. For the agricultural sector:

$$p_{a,2005}^{GGDC} = p_{1,2005}^{GGDC}.$$

For the non-agricultural sector:

$$p_{na,2005}^{GGDC} = \frac{\text{National Price Non-ag Value-added in Year 2005}}{\sum_{i=2}^{10}\left(\text{National Price Value-added of Sector i and Year 2005}/p_{i,2005}^{GGDC}\right)}.$$

These measures are equivalent to the prices relative to the price of US GDP in 2005:

$$p_{j,2005}^{GGDC} = \frac{\text{Price of sector } j \text{ output in 2005 US dollar}}{\text{Price of US GDP in 2005}}, \; j = a, na.$$

Secondly, since our expenditure data is indexed to the price level of urban Beijing in 2003, we need to convert $p_{j,2005}^{GGDC}$ accordingly.

Denote $\kappa_t^{BJ} = \frac{p_{c,t}}{p_{c,2003}^{BJ}}$ as the price of a reference consumption basket in year t relative to the 2003 price of a consumption basket in urban Beijing, where $p_{c,t}$ is in current RMB.[12]

[12]The effective expenditure is then $e = \frac{\text{Income in year } t \text{ RMB}}{\kappa_t^{BJ}}$.

Let $\Omega_{j,2005} = p^{GGDC}_{j,2005} \times$ Price of US GDP in 2005 $\times E_{2005}$, where $\Omega_{j,2005}$ is the price of sector js output in 2005 RMB. Using the sector-specific GDP deflator from the NBS with the year 2005 normalized to one, we define $\Omega_{j,t} = \Omega_{j,2005} \times \text{deflator}_{j,t}$ as the price of sector js output in current RMB. Finally, we convert the prices into 2003 Beijing RMB as follows:

$$p_{j,t} = \Omega_{j,t}/\kappa^{BJ}_t,\ j = a, na.$$

Thus, our sector prices are:

$$p_{j,t} = \frac{p^{GGDC}_{j,2005} \times \text{ Price of US GDP in 2005} \times E_{2005} \times \text{deflator}_{j,t}}{\kappa^{BJ}_t}.$$

Since we do not know $\bar{p} =$ Price of US GDP in 2005 x E_{2005}, we infer it from the data, and we will discuss how we calibrate $\bar{p}$ in the next section. Hence,

$$p_{j,t} = P_{j,t}\bar{p},$$

where $P_{j,t} = \frac{P^{GGDC}_{j,2005} \times \text{deflator}_{j,t}}{\kappa^{BJ}_t}$. The calibrated values of $P_{j,t}$ are shown in Table F.1.

The iceberg distributional costs, κ, in the urban and rural areas are measured by the consumption price data posted by Carsten Holz on his webpage. We update the original series reported in Brandt and Holz (2006) to more recent years. These prices vary across rural/urban regions, provinces, and years.

F.2. *Calibrating Consumption Preference Parameters*

Since the non-homothetic CES utility does not have the aggregation property, we need to use household expenditure data to calibrate the preference parameters. Consider the households with young agents and no migration. Then, all their income can be deflated by the rural distribution cost, κ_r. For each household in the group, the measured agricultural expenditure share in the model is:

$$\frac{e_a}{e} = \frac{\kappa_r p_a c_a}{\kappa_r e} = \frac{p_a c_a}{e} = \frac{\varphi_a p_a^{1-\varepsilon} c^{1-\varepsilon} e^{\varepsilon}}{e} = \varphi_a p_a^{1-\varepsilon} c^{1-\varepsilon} e^{\varepsilon-1}$$

where $c = c(e)$ as determined by the non-homothetic CES utility function. We use non-linear least squares to estimate $\varphi_a, \varepsilon, \mu$ and $\bar{p}$ that minimize the following objective function:

$$\sum_{\text{families with no elderly nor migrants}} \left[\frac{e_a}{e} - \frac{\widetilde{e_a}}{e} \right]^2 .$$

where $\frac{\widetilde{e_a}}{e}$ is the agricultural expenditure share observed in the data. We choose to match the agricultural expenditure share for families without elderly members or migrants because their expenditures are limited to rural areas and are thus more homogeneous.

The calibration results are shown in Table 4. The elasticity of substitution between agricultural and non-agricultural consumption (ε) is 0.342 and the income elasticity of non-agricultural goods (μ) is 2.446. The income elasticity is smaller for agricultural goods than for non-agricultural goods, suggesting that relative demand for agricultural goods declines with income. The calibrated $\bar{p} = 16.312$.

TABLE F.1

CALIBRATION RESULTS: SECTOR PRICES

year	P_a	P_{na}
2003	0.341	0.319
2004	0.392	0.335
2005	0.384	0.346
2006	0.388	0.359
2007	0.435	0.377
2008	0.461	0.386
2009	0.466	0.392
2010	0.499	0.409
2011	0.529	0.418
2012	0.537	0.414
2013	0.541	0.409

REFERENCES

Adamopoulos, T., L. Brandt, J. Leight, and D. Restuccia (2022). Misallocation, selection and productivity: A quantitative analysis with panel data from China. *Econometrica 90*(3), 1261–1282. [19]

Alvarez, J. (2020). The agricultural wage gap: Evidence from Brazilian micro-data. *American Economic Journal: Macroeconomics 12*(1), 153–173. [19, 23]

Angrist, J. D. and J.-S. Pischke (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press. [21]

Brandt, L. and C. A. Holz (2006). Spatial price differences in China: Estimates and implications. *Economic Development and Cultural Change 55*(1), 43–86. [8]

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica 69*(5), 1127–1160. [16, 19, 20, 26]

Cornelissen, T., C. Dustmann, A. Raute, and U. Schönberg (2016). From LATE to MTE: Alternative methods for the evaluation of policy interventions. *Labour Economics 41*, 47–60. [16, 19]

Eisenhauer, P., J. J. Heckman, and E. Vytlacil (2015). The generalized roy model and the cost-benefit analysis of social programs. *Journal of Political Economy 123*(2), 413–443. [19]

Fan, J. (2019). Internal geography, labor mobility, and the distributional impacts of trade. *American Economic Journal: Macroeconomics 11*(3), 252–88. [2]

Gollin, D., D. Lagakos, and M. E. Waugh (2014). Agricultural productivity differences across countries. *American Economic Review 104*(5), 165–70. [17]

Hamory, J., M. Kleemans, N. Y. Li, and E. Miguel (2021). Reevaluating agricultural productivity gaps with longitudinal microdata. *Journal of the European Economic Association*. [19, 23]

Hao, T., R. Sun, T. Tombe, and X. Zhu (2020). The effect of migration policy on growth, structural change, and regional inequality in China. *Journal of Monetary Economics 113*, 112–134. [19]

Heckman, J. J. and B. E. Honore (1990). The empirical content of the Roy model. *Econometrica 58*(5), 1121–1149. [16, 18, 19]

Herrendorf, B. and T. Schoellman (2018). Wages, human capital, and barriers to structural transformation. *American Economic Journal: Macroeconomics 10*(2), 1–23. [19, 23]

Lagakos, D., S. Marshall, M. Mobarak, C. Vernot, and M. E. Waugh (2020). Migration costs and observational returns to migration in the developing world. *Journal of Monetary Economics 113*, 138–154. [7, 16, 19, 23]

Lagakos, D. and M. E. Waugh (2013). Selection, agriculture, and cross-country productivity differences. *American Economic Review 103*(2), 948–80. [19]

Pulido, J. and T. Święcki (2018). Barriers to mobility or sorting? Sources and aggregate implications of income gaps across sectors and locations in indonesia. [16, 19, 23]

Schoellman, T. (2020). Comment on migration costs and observational returns to migration in the developing world. *Journal of Monetary Economics 113*, 155–157. [19, 22]

Tombe, T. and X. Zhu (2019). Trade, migration, and productivity: A quantitative analysis of China. *American Economic Review 109*(5), 1843–72. [19]

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources 50*(2), 420–445. [26]