THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF ECONOMICS

HUKOU AND GAOKAO: THE EFFECT OF THE HOUSEHOLD REGISTRATION SYSTEM AND REGIONALLY BIASED UNIVERSITY ADMISSIONS POLICIES ON LABOR MOBILITY AND HUMAN CAPITAL ACCUMULATION IN CHINA

DAVID STACK
SPRING 2016

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degrees
in Economics, International Politics, and Chinese
with honors in Economics

Reviewed and approved* by the following:

Russell Cooper
Professor of Economics
Thesis Supervisor

Russell Chuderewicz
Senior Lecturer, Department of Economics
Honors Adviser

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

This article fills a gap in the literature by examining the effect of China's *hukou* (household registration) system in the same context as the country's regionally biased university admissions policies. A two-period model is constructed to represent the migration and education decisions available to Chinese from rural areas, which is then calibrated using data from China Family Panel Studies, a database administered by the Institute of Social Science Survey at Peking University. After parameters defining the distribution of ability and "cost of moving" are estimated, comparative statics are performed to predict the effects of possible reforms to these policies. The results indicate that these policies act as serious impediments to both labor mobility and the efficiency of human capital accumulation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would first and foremost like to express my heart-felt gratitude to my thesis supervisor, Dr. Russell Cooper, for his tireless efforts to guide me through this process. I would be remiss, however, if I did not also take this opportunity to thank Dr. James Tybout for his helpful suggestions throughout the year. Finally, I would like to thank the Institute of Social Science Survey at Peking University for allowing me to access to the CFPS dataset.

Chapter 1

Introduction

One of the major conclusions of recent research on the causes of productivity differences across countries is that the systemic misallocation of resources on a micro-level can, in aggregate, lead to massive inefficiencies on the macro-scale. This misallocation can be induced by a variety of distortions, including limited spatial labor mobility and barriers that limit the ability of particular groups to acquire human capital or enter certain professions.[1] Two policies specific to China – one a relic of the Mao-era command economy, the other a product of the country's *de facto* federalist political system and highly unequal economic development – do precisely those two things, and in doing so act to retard the growth potential of the world's second largest economy.

The first is the *hukou* or "household registration" system, which limits labor reallocation across administrative boundaries and, more generally, from the rural sector to the urban sector. More specifically, this policy assigns every resident of China two attributes at birth: a "*hukou* type" (either agricultural or non-agricultural) and a "*hukou* location" (which corresponds with a particular village or urban residential community). The combination of these two attributes, which are inherited from one's parents, has an outsize impact on the educational and occupational opportunities afforded to each individual. For many years, migration from rural to urban areas was strictly prohibited. In the contemporary reform era that began in 1978, this restriction on movement has been gradually relaxed, to the point that today there are between

[1] For an excellent exploration of the latter, see Hsieh, Hurst, Jones, and Klenow (2013).

160 and 260 million agricultural, non-local *hukou* holders living in officially-defined urban areas. While these migrants are permitted to work in China's cities, they are typically relegated to the informal sector of the labor market working without a legally binding contract, and are thus particularly prone to exploitative employment practices. Furthermore, they are typically excluded from locally provided social services like education and healthcare, or they are given separate services of substantially diminished quality. There are now various policies in place that allow some rural-*hukou* holders to be "naturalized" into their city of residence and acquire urban-*hukou* status, but these programs vary widely from city to city, and it is generally much harder to officially migrate to a "tier one" city like Beijing or Shanghai than so-called "lower tier" cities in China's vast interior.

The second policy in question is the system through which college-bound Chinese citizens are assigned to universities.[2] Chinese high school seniors take what is known as the *gaokao* or National Higher Education Entrance Examination, which is analogous in many ways to the SAT in the U.S., except that an individual's *gaokao* score is effectively the sole criterion used for admission to a university. It is necessary to point out several important regionally biased facets of the Chinese admissions system. First, the *gaokao* is administered on a provincial basis and applicants are only compared to other test takers from their home province, as defined by their *hukou* registration (thus the child of two migrant workers in Beijing must return to her parents' home province to take the exam). Each university sets a quota for students from each province which corresponds less with the population of each province and more with the perceived quality of students from the province. Critically, universities set significantly higher

[2] The term *assigned* is employed deliberately to emphasize that, despite the fact that Chinese students are free to apply to any universities they so choose, the basic purpose of a university admissions system is to *match* students with universities based on any number of criteria.

admissions quotas for applicants from the province the university is located in, and more prosperous provinces tend to have significantly more top-tier higher education institutions, particularly on a per-capita basis.[3] Thus high ability students from less developed provinces are doubly penalized for coming from provinces with lower perceived student quality and not having the benefit of easier admission to multiple high-tier universities.

An additional nexus exists between these two policies in the sense that one of the most important pathways for rural-born Chinese to acquire urban *hukou* status in a major city is to obtain a high-skilled job in that city. This method of officially sanctioned migration requires a college degree, and it helps considerably if the university one attended is considered to be of high quality or if that university is situated in the same location of the sought-after job.

For these reasons, the research design of this study was selected in the belief that the *hukou* and *gaokao* policies are so intertwined that it is imprudent to examine either one in isolation. More specifically, I construct a two-period model of the migration and education choices available to rural Chinese citizens and then use data from the China Family Panel Studies dataset to calibrate that model and estimate its key parameters.[4] I then perform comparative statics in order to simulate the effects of potential changes to the two policies under consideration. The results of these policy experiments indicate that even fairly small changes to these policies could lead to large changes to the fractions of the rural-born population that

[3] By way of example, consider the case of Beijing and Henan province. Beijing has a population of 21.5 million, a per capita GDP of $17,000, and twenty-five universities in the officially defined list of the top 116 universities in China (Project 211). Henan has a population of 94.3 million, a per capita GDP of $6,300, and only one top-tier university.

[4] China Family Panel Studies (CFPS) is funded by the 985 Program of Peking University and carried out by the Institute of Social Science Survey of Peking University.

migrates to cities and/or enrolls in high-tier universities. Ultimately, such reforms have the potential to improve both the equity and efficiency of the Chinese economy.

The layout of this paper is as follows. Chapter 2 presents a review of the relevant literature. Chapter 3 describes the methodology used to construct and calibrate the model and perform “policy experiments,” while Chapter 4 elaborates on the key data sources, assumptions, and modeling decisions. Chapter 5 contains a summary of the inputs in the baseline model, the results of the policy experiments, and an interpretation of the findings. Chapter 6 concludes.

Chapter 2

Literature Review

2.1 Misallocation in China

The reform era that began in 1978 has been characterized by consistently high annual rates of economic growth, a trend that has been driven in no small part by increases in aggregate productivity of the Chinese economy, which has in turn been largely caused by reductions in misallocation of capital and especially labor. Brandt, Hsieh, and Zhu (2008) note that between 1978 and 2004, aggregate labor productivity doubled every ten years. The authors hypothesize that this impressive growth record was the result of two major structural shifts in the Chinese economy: reallocating labor from the agricultural sector to manufacturing and services, and reallocating labor and other resources from state-owned enterprises (SOEs) to nonstate firms. They create a calibrated, dynamic three-sector model of the Chinese economy that they then use to run a series of counterfactual experiments in order to try to isolate the causal impact of each shift.

The authors point out that in pre-reform China two highly unproductive sectors of the economy – the collectivized agricultural sector and the state-owned industrial sector – dominated employment. Despite a "6:1 gap in output per worker between the nonagricultural and the farm sector," a full 70 percent of Chinese were employed in agriculture in 1978. Within the urban sector that year, SOEs employed 80% of the labor force. This situation, a clear case of misallocation, made the country particularly well suited for taking advantage of the gains from labor reallocation.

Brandt, Hsieh, and Zhu find that these two structural transformations in the Chinese labor market both had significant effects in driving growth, but the transfer or workers into the nonstate sector had a more significant effect than the transfer of workers into the nonagricultural sector. More specifically, these transformations contributed 2.29 percentage points to China's overall growth rate during this time period, or about one third of the total. "Out of this 2.29 percentage point contribution," they write, "the reduction in labor market barriers between the state and nonstate sector accounts for 1.77 percentage points, while the reduction in barriers between agriculture and nonagriculture accounts for only 0.52 percentage points." Ultimately, the authors conclude that the most source of China's growth in the reform period has been TFP growth in the nonstate sector.

Brandt, Tombe, and Zhu (2013) examine misallocation in China's nonagricultural economy, with a focus on the spatial dimension. The authors make an important contribution by taking the basic methodology of Hsieh and Klenow (2009) and applying it on a more aggregated, and thus perhaps more informative, scale. They rightly point out the high degree of "local autarky" exhibited by China's administrative units, particularly provinces, prior to reform. They explain that "self-sufficiency in both agriculture and industry were aggressively pursued, and reinforced through limited investment in transportation infrastructure." The net effect was that the inter-regional flow of labor, goods, and capital was severely restricted. Reform has meant the gradual elimination of these restrictions; the authors attempt to determine the net effect of the resulting reallocation across administrative boundaries. Furthermore, while Hsieh and Klenow (2009) analyze misallocation *within* sectors, Brandt, Tombe, and Zhu (2013) examine distortions affecting the distribution of factors *across* sectors.

Despite the inclusion of between-province labor reallocation within their analytical framework, the authors' ultimate conclusion is broadly similar to that of Brandt, Hsieh, and Zhu (2008), i.e. misallocation between the state and nonstate sectors represents the most important restraint on TFP. The authors find that "this distortion accounts for most of the within-province distortions and, more importantly, almost all the time variation in the impact of distortions."

At the same time, Brandt, Tombe, and Zhu still have important findings for those seeking to analyze the misallocation of labor across space in China. Interestingly, the authors find that despite the massive interprovincial migration that has occurred in recent decades, the amount of TFP reduction associated with between-province labor market barriers has remained essentially constant over time. This is surprising because, *ceteris paribus*, migration would decrease in the difference in provincial returns to labor. The authors offer the following explanation: the reallocation of labor has been "not fast enough to offset the rising dispersion of TFP" resulting from the economic boom in the coastal provinces.

2.2 The Effect of the Hukou Policy on Labor Market Efficiency

There is no shortage of economic research on the effects of the hukou policy on the segmentation of Chinese labor markets or on the welfare of rural-urban migrants themselves. With regards to the former, many scholars point out that there is a relatively sharp divide in urban labor markets between jobs for migrants and jobs for urban natives. Herd, Koen, and Reutersward (2010) provide detailed data on this phenomenon drawing from the 2005 1% Census data. That year, 52% of urban natives were employed in the public sector (either for the government directly or a SOE), whereas that figure is only 5% for unofficial rural migrants. While 42% of local urban hukou holders had white-collar jobs in 2005, only 7% of migrants did. This segmentation is tangible enough for Knight, Song, and Jia (1999) to conclude that "urban and rural migrant workers are not close substitutes in the production function of urban firms."

There is also ample evidence of the discrimination faced by rural migrants, and how this effects their economic and social wellbeing. Meng and Zhang (2001) find that the difference in wages between migrants and urban natives cannot be wholly explained by differences in productivity-related differences. While largely pursuing different occupational paths, when migrants and natives work the same job, natives often make significantly more. In fact, the authors find that "82% of the hourly wage differential between the rural migrants and urban residents is due to intra-occupational wage differentials." Summarizing the economic situation migrants find themselves in, Akgüç, Liu, and Tani (2014) write, "migrants work more hours, have lower wages, jobs of worse quality and tenure than" urban natives.

Research on how the hukou policy impacts the allocative efficiency of the Chinese economy is considerably less abundant, but certainly exists. Yang and Zhou (1999) indicate that in 1992, massive disparities existed in the marginal product of labor across three sectors: the state industrial sector (9,346 yuan), the rural industrial sector (1,211 yuan), and agriculture (601 yuan). Convergence does not occur due to insufficient migration, which the authors attribute to a combination of how non-*hukou* migrants can't access urban social services, the lack of an effective land tenure system, and the high cost of child care and schooling in cities. Au and Henderson (2006) employ standard models of urbanization to contend that China's cities are substantially smaller than they would be in the absence of migration restrictions, a situation that leads to "under agglomeration" on a scale that they claim is "enormous." Ito (2008) builds upon the foundations of the Harris-Todaro model of migration, but updates the simple dual economy model to reflect seven different regions and four different sectors. He runs a simulation-style analysis, and finds that, if the hukou system were to be eliminated, it would result in massive migration to the cities (to the detriment of the rural industrial sector). Importantly, he also finds that this would only lead to large-scale reductions in income inequality if the rural population increases its average education level and if baseless discrimination against those from rural areas would cease to exist

in urban labor markets.

2.3 Hukou in the Context of Higher Education

While there is a voluminous literature on Chinese higher education in general and the heterogeneity of access and returns to higher education more specifically, there are considerably fewer studies that examine both higher education and *hukou* in the same context. Wang and Moffatt (2007) conduct their own survey of recent graduates from three universities in Wuhan. After examining the data, the authors find evidence that, due to the increased pressure to secure an urban job, college graduates with rural *hukou* invest in higher job-search effort, set lower reservation wages, and accept lower starting salaries than their urban counterparts.

Liao et al. (2015) perform the study that is the most comparable to this one in terms of both research topic and design. Just like this study, the authors conduct a model of migration and education decisions, which they calibrate with real data, and then run policy experiments. There are several key differences between their study and this one, however. In some ways, Liao et al.'s model is more realistic than this one. First and foremost, Liao et al. construct a model that incorporates dynamic equilibria into its design, while this paper proposes a model of individual choice. Furthermore, Liao et al. attempt to decompose the contribution of the various methods of migration on overall economic growth. That is outside the research scope of this article.

In other ways, however, this study is richer and more predictive. First, while in Liao et al.'s model individuals vary only in ability, in this paper individuals are heterogeneous in two dimensions: ability and cost of moving. Next, Liao et al. set many of their key parameters using figures from the past work of other researchers and sometimes other countries (for example their ability distribution parameter is based on the work of Feenberg and Poterba [1993], who estimated a shape parameter for income inequality in the United States). In this study, parameter estimates are based on same data that the other model inputs are drawn from. Finally, Liao et al.'s policy experiments attempt to answer the questions of what would have happened if policies like graduate job assignment or strict migration restrictions had remained in place. This paper is more forward-looking in its approach. It attempts to predict the effect of what will happen to the choices of rural residents if economic liberalization continues.

Chapter 3

Methodology

3.1 Modeling the Choices of Rural Residents

The following two-period model is constructed in order to the represent the choices faced by newly college-aged Chinese citizens from rural areas and low-tier cities. Individuals are heterogeneous in their natural ability (θ) and the cost of moving (M) that they would incur by leaving their officially designated residence. This cost of moving is defined as the sum of financial, psychological, and policy-induced costs necessary to move one's person and possessions to a new city. θ and M are assumed to be statistically independent and distributed along exponential distributions defined by rate parameter α in the case of θ and parameter λ for the distribution of M.[5] Individuals live for two periods of length t. They maximize utility by making a discrete choice between eight explicitly defined life paths, each of which is associated with payoff function defined in terms of the discounted present value of lifetime income (Y).

In period 1, individuals must decide whether to enroll in a university or immediately enter the full-time labor force. In both cases, rural residents also choose whether to engage in these activities in their home region or in a representative "major city." Those who choose to not enroll in a university earn an exogenously defined wage for unskilled labor in the rural sector in period 1 (ω_1^{UR}), while those who enter the urban sector earn can earn a higher wage (ω_1^{UU}) with probability π^{UR} (i.e. the probability that an unskilled worker with rural hukou can get a job in the urban sector and their wages will not be arbitrarily withheld). If migrants cannot secure a job, they earn an income of zero for that period, but they can try to secure employment again in either

[5] The assumption that M and θ are distributed along exponential distributions was made solely to simplify the process of parameter estimation.

sector in the following period. Importantly, migrants also incur their personal cost of moving (*M*). Individuals who choose to enroll in a "low-tier university" (located in rural areas and low-tier cities)[6] can accumulate human capital worth h_L, which is less valuable than what they can accumulate at a university located in a major city (h_H). During this period, students spend an amount of time ($\bar{e}$) in school, and spend the remainder of their time working at the prevailing skilled wage rate for those in their location and age group (ω_1^{ER} or ω_1^{EU} for the rural and urban sectors, respectively). They must also pay a tuition rate (P_L or P_H depending on the quality of the university) multiplied by the time spent in school, which is fixed at $\bar{e}$. Finally, just as for non-students, individuals who choose to enroll in a university in a major city must incur a cost of *M*.

In period 2, higher education is no longer an option. Individuals take the human capital they may have accumulated in period 1 and decide whether to work in the rural or urban sector. Unskilled workers earn ω_2^{UR} in the countryside and ω_2^{UU} in the city. The probability of earning the unskilled urban wage depends on whether the individual in question has previously worked in the city before. Those who haven't face the same probability as similar individuals from period 1 (π^{UR}), while unskilled workers who have already moved to the city face a probability of π^{UU}. Individuals who chose to receive in higher education during period 1 earn a wage equal to their natural ability (θ) multiplied by their human capital multiplied by the wage for educated workers in the location of their choosing (ω_2^{ER} or ω_2^{EU} for the rural and urban sectors, respectively). Educated workers searching for a job in the urban sector secure one with probability π^{EU} if they were educated in at a high-tier university and with π^{ER} if they studied at a low-tier university. Importantly, for all workers in period 2, both skilled and unskilled, moving to the city entails a cost of *M* while the reverse process is costless. Finally, all workers evaluate

[6] Very few Chinese universities are actually located in officially defined *rural* areas. High-prestige universities are disproportionately located in provincial capitals.

earnings from period 2 by dividing by an interest rate R (all individuals are homogeneous in their level of patience). Table 1 summarizes these definitions:

Table 1: Variable Definitions

Variable	Definition	Variable	Definition
θ	Individual's ability	ω_1^{UR}	Wage for *unskilled* workers in the *rural sector* in *period 1*
M	Individual's cost of moving	ω_1^{UU}	Wage for *unskilled* workers in the *urban sector* in *period 1*
α	Rate parameter for ability	ω_2^{UR}	Wage for *unskilled* workers in the *rural sector* in *period 2*
λ	Rate parameter for cost of moving	ω_2^{UU}	Wage for *unskilled* workers in the *urban sector* in *period 2*
Y	DPV of lifetime income	ω_2^{ER}	Wage for *educated* workers in the *rural sector* in *period 2*
π^{UR}	Probability of securing an *unskilled* job in the urban sector after moving from the *rural* sector	ω_2^{EU}	Wage for *educated* workers in the *urban sector* in *period 2*
h_L, h_H	Human capital from attending a *low-* or *high-tier* university	t	Years in each period
$\bar{e}$	Proportion of period 1 spent in university	R	Interest rate

These choices are summarized visually in Fig. 1 with a decision tree representing all eight possible life paths and the payoff function associated with each choice:

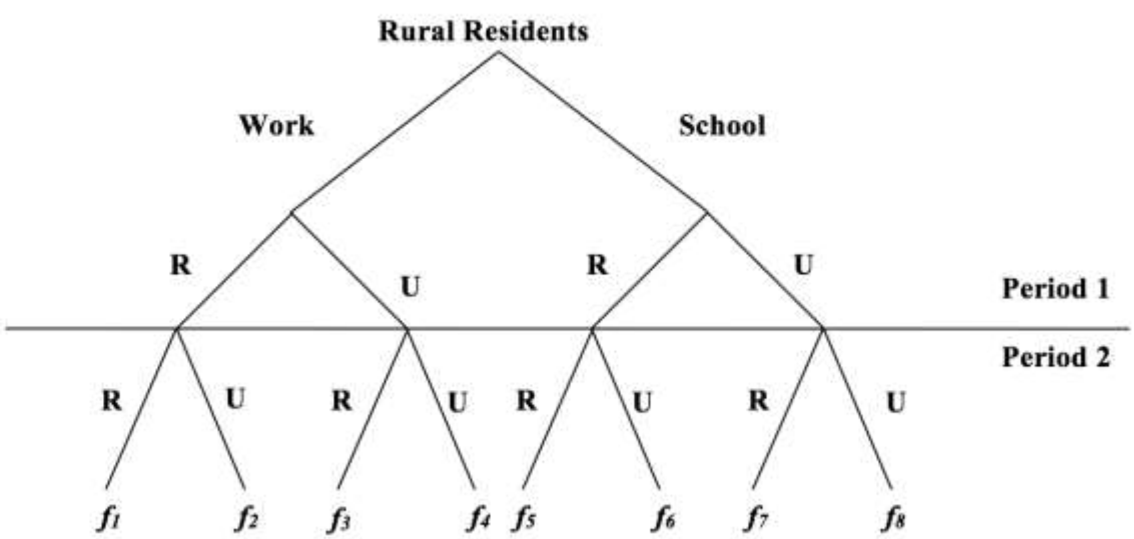

Figure 1: Possible life paths for rural residents

The payoff functions for each life path are as follows:

$$\boldsymbol{f_1} = \omega_1^{UR} + \frac{\omega_2^{UR}}{R}$$
$$\boldsymbol{f_2} = \omega_1^{UR} + \frac{\pi^{UR}[\omega_2^{UU} - M]}{R}$$
$$\boldsymbol{f_3} = \pi^{UR}[\omega_1^{UU} - M] + \frac{\omega_2^{UR}}{R}$$
$$\boldsymbol{f_4} = \pi^{UR}[\omega_1^{UU} - M] + \frac{\pi^{UU}[\omega_2^{UU}]}{R}$$
$$\boldsymbol{f_5} = (1 - \bar{e})(\omega_1^{ER}) - P_L\bar{e} + \frac{\theta h_L \omega_2^{ER}}{R}$$
$$\boldsymbol{f_6} = (1 - \bar{e})(\omega_1^{ER}) - P_L\bar{e} + \frac{\pi^{ER}[\theta h_L \omega_2^{EU} - M]}{R}$$
$$\boldsymbol{f_7} = \pi^{UR}(1 - \bar{e})(\omega_1^{EU}) - P_H\bar{e} - M + \frac{\theta h_H \omega_2^{ER}}{R}$$
$$\boldsymbol{f_8} = \pi^{UR}(1 - \bar{e})(\omega_1^{EU}) - P_H\bar{e} - M + \frac{\pi^{EU}[\theta h_H \omega_2^{EU}]}{R}$$

Given the inputs on the baseline model (described later), three payoff functions are strictly dominated: f_2, f_3, and f_7, i.e. the functions corresponding with life paths in which an individual switches sectors after period 1 (except for f_6, in which individuals attend low-tier universities and then migrate to high-tier cities). As a result, when all eight functions are plotted in a three-dimensional space (with M-, θ-, and Y-axes), a vertical perspective shows only the five functions chosen by rational individuals given their cost of moving and ability level. This is because individuals maximize lifetime income by selecting the life path that corresponds with the highest plane above their endowment of M and θ. This is represented visually below using the "baseline" model inputs (described later) and an artificial domain cutoff of (0,1) for both M and θ for easier viewing:

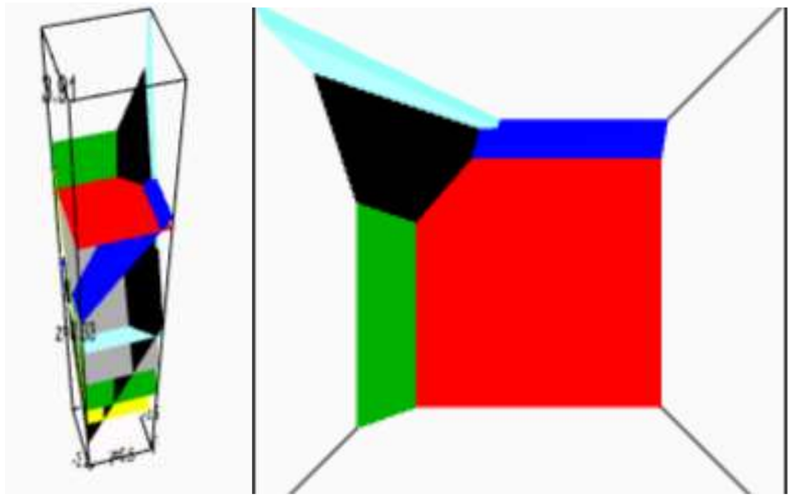

Figure 2: Graph of payoff functions from a side angle and a vertical perspective

In Fig. 2, the x-axis is M, the y-axis is θ, and the z-axis (only visible in the left figure) is Y. The top-left (turquoise) section corresponds with f_8, i.e. enrolling in an urban university and then working in an urban area. As the graph shows, this life path is chosen by individuals with high θ but low M. f_8 is subject to an exogenous θ cutoff, below which it is equal to 0 (explained later). Moving clockwise, the blue section represents f_5 (enrolling in a low-tier university and working in the rural sector), red is f_1 ("staying on the farm"), green is f_4 (becoming a low-skilled migrant worker), and the black section is f_6 (attending a low-tier university but then seeking employment in a major city in period 2). This shows that, under certain conditions, the model predicts that high ability individuals will enroll in universities, and individuals with high moving costs will stay closer to home.

The following methodology is employed to find the percentage of rural residents selecting each life path. By setting two payoff functions equal to each other, one can solve for the line along which individuals are indifferent between the two options. Doing this for every possible pair of the five functions shown on the right side of Fig. 2 yields the seven lines that define the boundaries of the regions shown in that figure. These boundary functions can then be used to find the volume under the joint probability distribution function, which is defined as:

$$f(M, \theta) = \lambda \alpha e^{-(\lambda M + \alpha \theta)}$$

This function is represented graphically below (using arbitrarily chosen parameters):

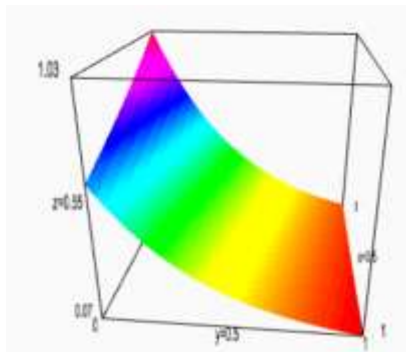

Figure 3: Joint Probability Distribution Function

Note that the figure above indicates that individuals are much more likely to have low M and low θ than high M and θ. Taking the double integral of this PDF over a region defined by the boundary functions found earlier can give the volume under any plane shown in Fig. 2, which is equal to the fraction of the population choosing that particular life path. Take, for example, the volume under the plane defined by f_4, i.e. the lower-left, green region of Fig. 2. To find the fraction of the population choosing to work in the urban sector in periods 1 and 2, the following integral is computed:

$$\int_0^{M^*} \int_0^{g(M)} \lambda \alpha e^{-(\lambda M + \alpha \theta)} \, d\theta \, dM$$

In this case M^* is a constant equal to the level of M at which individuals are indifferent between f_4 and f_1 and $g(M)$ is the boundary line separating f_4 and f_6 in Fig. 3. Due to the more irregular shape of the other regions in Fig. 3, some need to be split into sub-regions for the sake of easier integration, but the same basic principle applies.

In order to incorporate the effects of China's regionally-biased university admissions system into the model, an exogenous θ threshold is introduced. First, we assume that a student's score on China's national university admissions examination or *gaokao* can be modeled by the following:

$$score_i = f(\theta_i) + \varepsilon_i$$

Then we can introduce an artificial threshold of θ set at θ^T to model the way China's admissions system constructs a barrier that those from rural backgrounds seeking to enroll in top-tier universities must overcome. A higher *gaokao* "floor" for rural students, i.e. a higher θ^T, will thus force some students who would otherwise enroll in a high-tier university to enroll in a low-tier school. In this way, the model shows the potential for policy-induced educational mismatch.

Incorporating the effects of the *hukou* policy works a bit differently. As mentioned before, the cost of moving, M, is defined as the sum of "natural" moving costs, like the financial cost of transporting one's person and possessions to a new city along with psychological cost of doing so, and *hukou*-induced costs. By altering the distribution parameter of M, i.e. λ, such that the average cost of moving falls by a particular proportion, it is possible to simulate the effects of a reform to the *hukou* policy that lowers M by the same amount.

3.2 Estimating Parameters by Matching Moments

The model's inputs are the appropriate numerical values for exogenous variables (like wage levels) and the parameters α and λ, which define the distributions of θ and M, respectively. Given these inputs, the model is able to predict the proportion of the population choosing each life path. Minimizing the squared distance between the model-predicted fractions of the rural population choosing each life path and the actual figures, as represented in the data, by solving for optimal values of α and λ yields an estimate of these parameters. This can be presented more formally with the following multivariable optimization problem:

$$\min_{(\alpha,\lambda)} (m^S - m^D)^2$$

In the objective function above, m^S and m^D are (5 x 1) matrices representing the population fractions or "moments" predicted by the model and found in the data, respectively.

Although simple in theoretical structure, the complex nature of each component of m^S makes solving this optimization problem computationally difficult in practice. Each component, which corresponds with the model predicted fraction of the population choosing a particular life path, is a double integral of the joint probability distribution function defined over a typically non-rectangular region. For this reason, minimizing the objective function using partial derivatives is not only impractical to perform by hand but also too computationally intense for the primary computational mathematics software used by the author, i.e. SageMathCloud. That same software package, however, makes it easy to employ the "limited-memory Broyden-Fletcher-Goldfarb-Shanno" optimization algorithm (or "L-BFGS-B," with the final "B" referring to the bounded constraints on possible solutions). A further complication arises from the shape of the optimization function in question, which is characterized by a long, narrow, and parabolic trough or valley similar to that of the Rosenbrock function.[7]

[7] The Rosenbrock function is commonly used to test the efficacy of optimization algorithms.

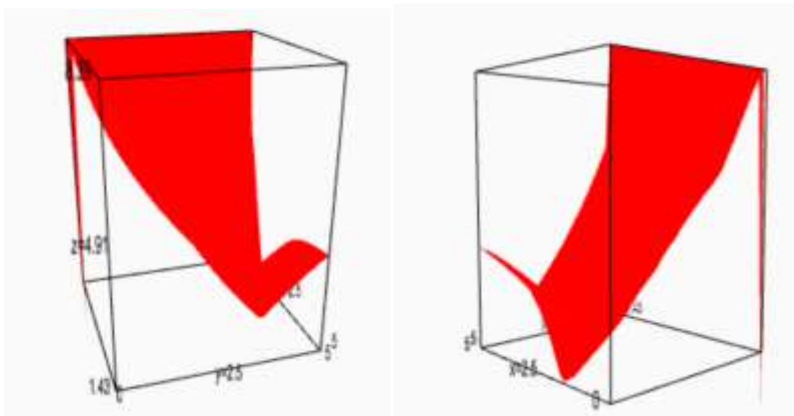

Figure 4: An exaggerated form of an example objective function from alternate perspectives

Due to this shape, the output of the L-BFGS-B optimization algorithm is highly sensitive to the "first guess" inputted by user. It is possible, however, to use every possible combination of integer values of α and λ between 0 and some reasonable, yet arbitrary, maximum, for example 10. For some combinations of model inputs (see Chapter 4 for more discussion on model variation), there are no issues, and all possible first guesses yield roughly the same result, but sometimes a situation results in which many resulting estimates of α and λ are nowhere close to each other. When this occurs is usually the case that the algorithm outputs lie along the parabolic trough shown above (See Fig. 5 below for a visual representation).

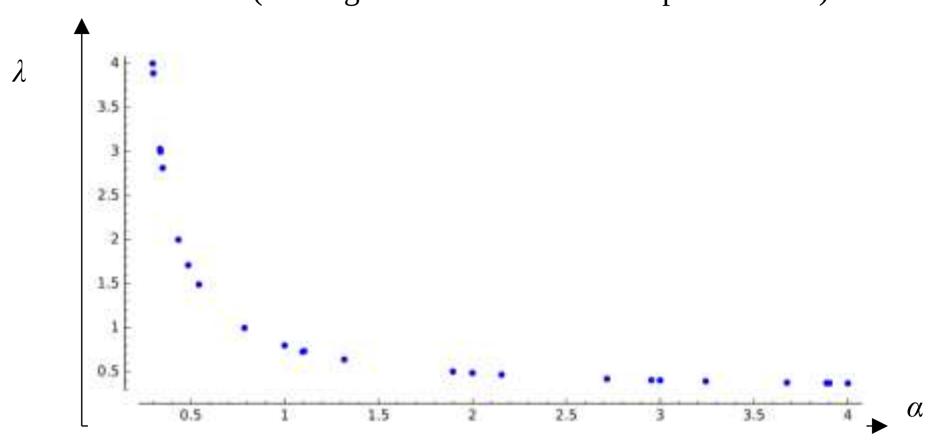

Figure 5: An example of a less-than-ideal algorithm output (points represent potential minima)

Even when this occurs, however, it is possible to test every point produced by the algorithm to determine the local minimum within a reasonably defined domain of possible parameter values. Fortunately, in most cases there is clear clumping around what further testing reveals to be the local minimum (See Fig. 6):

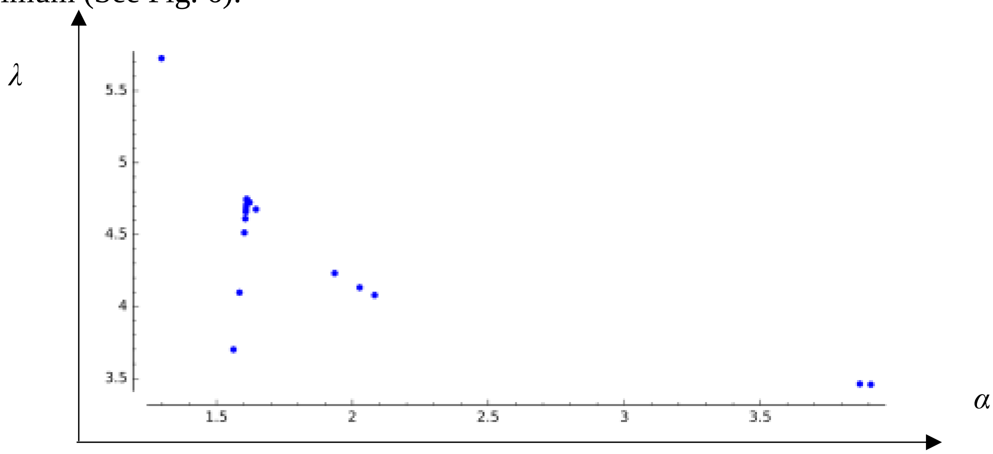

Figure 6: A more typical example (note the clumping around [1.6,4.7])

In this way, given a set of exogenous variable values and data-generated moments it is possible to estimate the distribution parameters of M and θ with reasonable accuracy.

3.3 Comparative Statics and Policy Experiments

At this point, under the assumption that these parameters are invariant, i.e. they represent fundamental structural characteristics of the population that do not change in response to exogenous shocks, it is possible to perform policy experiments to simulate the distribution of the rural population across life paths under different circumstances and to help predict the way the Chinese economy would respond to changes to the *hukou* and *gaokao* policies. For example, it is possible to calculate how the magnitude of the decline in average moving costs necessary to induce a particular fraction of those from rural areas to leave the countryside. Alternatively, one could test how life path distributions would react to a decrease in the binding *gaokao* floor rural-dwellers need to surpass to enter a prestigious university.

Chapter 4

Data Sources, Assumptions, and Key Modeling Decisions

4.1 Data Sources

The primary source used for information on individual-level characteristics and decisions is the China Family Panel Studies (CFPS) dataset. This source contains information on 1,744 variables for 35,179 individuals from all of China's provincial level administrative divisions except for Inner Mongolia, Tibet, and Hainan (although provinces are not represented in the survey in accordance with their share of the Chinese population, and sample sizes for some regions – like Qinghai and Ningxia – are extremely small). Response rates for any given question vary widely, and the average response rate is quite low. Each individual is attributed with a six-digit "community ID," although these ID numbers do not correspond with the widely used six-digit *Guobiao* (national standard) codes for county-level subdivisions, and the identity of each CFPS community ID is not released by ISSS in order to preserve the privacy of respondents. It is, however, possible to know whether each individual was surveyed in an "urban" community (according to three different definitions of that term) and which province the survey was taken in. Other relevant variables include income, age, current hukou status (agricultural or non-agricultural), current province of hukou registration, level of education, location of *currently* attended school (options include provincial capital cities, "general cities," county seats, and villages; unfortunately, there is no variable for location or quality of school attended in the *past*), educational expenses, the year the individual left their place of registration, whether current the individual's current residence matches his or her official *hukou* location at age 12, and whether the individual felt discriminated against due to his or her *hukou* status. Other data sources include the World Bank data on real interest rates and figures related to wage differentials between various categories of educated workers from Zhong (2011) and Wang and Moffatt (2008).

4.2 Core Assumptions

The construction of this model and simulating policy changes by altering its exogenous variables relies on a number of a core assumptions that deserve scrutiny. First, and perhaps most critically, is the assumption that the set of exogenous variable inputs and moments from the data used to "calibrate" the model are roughly accurate representations of 1) the relative utility of the eight different life choices available to rural individuals given their ability and cost of moving and 2) an accurate representation of the number of rural individuals who have selected each life choice. Whether this assumption is valid is somewhat dubious for a number of reasons, including the leaps of faith necessary to calculate the moments from the data due to the lack of accurate information on migrant status and location of completed higher education, the rapidly changing institutional environment (including the continuous reforms to the *hukou* system that have taken place during the period analyzed and the rapid expansion of access to higher education during that same time), and the fact that the exogenous variables in question have all changed considerably during this period. Manufacturing wages, for example, increased on average by 12% every year from 2001-2015 (The Economist, 2015). An additional, more minor assumption is setting the level of urban to rural migration to zero (this has especially been inaccurate in recent years). For these and other reasons, the results of this work should be met with a healthy level of skepticism. At the same time, while the robustness of the outputs is likely questionable, the methods used to find them should help to advance the literature on the subject.

4.3 Key Modeling Decisions and Variable Operationalization

Due to the lack of reliable information on a number of important values, several important modeling decisions need to be made. The model is sensitive to small changes in its inputs (if the changes are too large, entire life paths can disappear from the analysis by becoming strictly dominated), so it is critical that these are made with two factors in mind: the empirical validity of the choice and how it affects the model.

One extremely important decision is how to define “rural” and “urban.” For one thing, the administrative boundaries of Chinese cities often extend far into the countryside surrounding an urban area. Furthermore, even if one could perfectly isolate the *de facto* urban from the *de jure*, including the “urban core” of a “low tier” prefectural-level city like Lijiang in Yunnan province as an option for rural residents to migrate to distorts the original intent of the model. One could likely safely assume that outside of psychological and policy-induced costs, few rural-to-urban migrants would willingly choose to migrate to Lijiang as opposed to an East Coast economic powerhouse like Shenzhen. In order to address both of these difficulties, the author’s original research plan called for creating a dummy variable which captured *de facto* urban areas in Tier One and Tier Two cities, but the lack of geographic granularity in the CFPS data necessitated a different approach (described in the next chapter).

Next is the definition of the durations that demark periods 1 and 2. There is a major tradeoff at play in this choice. If period durations are defined too broadly, e.g. thirty years apiece, the moments found in the data would clearly not be representative of the actual choices undertaken by rural individuals due to China’s rapidly changing institutional environment and rapidly growing economy. In the most extreme case, individuals in the pre-reform era effectively made no economic choices whatsoever due to the structure of the Maoist economy. If durations are defined too narrowly, however, there can be substantial reductions in the sample size of the data, particularly for specific groups like educated, East Coast urbanites.

Furthermore, the data lacks clear differentiation between who is a migrant and who is not. For this reason, the choice regarding who to place in the category of a migrant or likely migrant has a major impact on the model results, as does whether (and how) to differentiate between the wages of migrants and non-migrants when calculating exogenous variables like wage levels.

Similarly, small changes in the definition of who is considered “educated” lead to wide variation in the moments found in the data. The two most obvious possible cutoffs are having completed a three-year degree program (which are typically more vocational in nature) or a traditional four-year bachelor degree. This is not a trivial choice. While 1,678 respondents reported having completed a three-year degree, only 1,016 in total said they had a four-year, master’s, or doctorate degree.

Chapter 5

Results and Discussion

5.1.1 Baseline Model Parameters and Elasticities

This subsection summarizes the modeling decisions and results of the most empirically defensible set of inputs attempted by the author.

The primary exogenous variables were drawn from the CFPS 2012 dataset. Period 1 individuals are defined as being between 18 and 26 years of age, while those in period 2 are between 27 and 35. The 35-year-old age maximum was selected because these individuals who were 35 in 2012 were 18 in 1995, the year university students started to apply for their own jobs after graduation as opposed to being assigned (Liao et al., 2015).

Urban status was defined using a dummy variable called “ECurborurb” that captures locations in the prosperous East Coast provinces of Beijing, Tianjin, Shanghai, Jiangsu, Zhejiang, Fujian, and Guangdong that are also *either* defined as “urban” by the census *or* by the CFPS survey taker (this “or” statement expands the pool of urban areas considerably).

Income outliers were removed by eliminating respondents who reported zero income or had incomes three standard deviations higher than the mean reported value for their particular category, e.g. educated urbanites. Furthermore, an attempt was made to capture the urban wage levels of migrants as opposed to the average of migrants and urban natives. For uneducated workers this was done by finding the average reported income of individuals located in an area that satisfies the requirements of “ECurborurb,” who had education levels below that of a four year degree, and who also satisfied at least one of a number of conditions meant to indirectly proxy for migrant status, including whether the individual’s *hukou* location didn’t match the province of the survey, whether the respondent reported that they felt discriminated against due to their *hukou* status, and whether they currently have agricultural *hukou*. None of these are perfect proxies, and the need for more than one stems from the low response rate for any given question. For educated workers, the low sample size for those in model-defined urban areas with four-year degrees or higher who also satisfy any proxy for being a migrant necessitated a different approach. Instead, the average income for those of migrant and non-migrant status was calculated and then discounted by 6%, which is based on the difference found between college graduates with urban and non-urban *hukou* by Wang and Moffatt (2008).[8]

The average real interest rate from 1995 – 2012 of 2.649% reported by the World Bank was raised to the ninth power, corresponding with the nine years between the midpoints of periods 1 and 2. All earnings in period 2 are discounted by being divided by the result, 1.265.

Tuition rates are differentiated between high- and low-quality schools by finding the average educational costs of those who reported attended universities in provincial capitals and those who didn't.[9]

Due to the lack of data on location or quality of *previously attended* universities, differentiating between the wages of educated migrants who attended high- and low-tier universities required turning to outside sources. More specifically, based on the findings of

---

[8] Wang and Moffatt find that annual income of recent graduates with non-urban *hukou* is 21,954 RMB, while those with urban *hukou* earn 24,981 RMB. It is possible to estimate the rural *hukou* wage from the average wage found in the data by setting up a system of two equations, $n = 1.138m$ and $.538m + .462n = w$, where m, n, and w represent the skilled wages of migrants, non-migrants, and the overall population, respectively and the coefficients represent fractions of the population that belong to each group in the sample (estimated using the method described below).

[9] All but four of the top 39 universities in China (as defined by the official "Project 985") are located in provincial capitals.

Zhong (2011) the author discounted the educated wages found previously, i.e. the f_8 wage level, by a further 14% to estimate the f_6 wage level.[10]

The "*gaokao* floor" was initially set arbitrarily to $\theta^T = 1.5$, which was later altered during the policy experiments. Determining how much of the population this floor allows into top schools requires knowledge of the parameter of that distribution, λ.

Finally, the exogenous inputs were normalized as proportions of a baseline wage level ω_1^{UR}. The baseline model exogenous inputs are summarized in Table 2 below:

Table 2: Baseline Model Exogenous Inputs

Variable	Value (Source)	Variable	Value (Source)	Variable	Value (Source)
$\boldsymbol{\omega_1^{UR}}$	1 (CFPS)	ω_2^{ER}	1.94 (CFPS)	π^{UR}, π^{UU}	.9 (assumption)
$\boldsymbol{\omega_2^{UR}}$	1.26 (CFPS)	$\omega_2^{EU} * h_H$	4.00 (CFPS, Wang and Moffatt)	π^{EU}, π^{ER}	1
$\boldsymbol{\omega_1^{UU}}$	1.27 (CFPS)	$\omega_2^{EU} * h_L$	3.44 (CFPS, Wang and Moffatt, Zhong)	θ^T	1.5
$\boldsymbol{\omega_2^{UU}}$	1.96 (CFPS)	ē	(4/9) = 0.444	t	9
$\boldsymbol{\omega_2^{UR}}$	1.96 (CFPS)	P_L	.527 (CFPS)		
$(\mathbf{1} - \text{ē})\boldsymbol{\omega_1^{ER}}$	0.88 (CFPS)	P_H	.776 (CFPS)		
$(\mathbf{1} - \text{ē})\boldsymbol{\omega_2^{EU}}$	1.10 (CFPS	R	1.265 (World Bank)		

[10] Zhong estimates the coefficients on dummies for attending a "very good" university and an "average" university are .54 and .362, respectively. It is then possible to employ a similar methodology to that of Footnote 8 to arrive at the 14% discount for attending a low-tier school.

Statistical moments were calculated using the 2012 CFPS data. Unfortunately, because there is no variable in the dataset on where a person's education took place or the quality of the school that an individual attended, only a general moment for people corresponding with both f6 and f8 could be computed. The author then split this overall number into two components weighted by the percent of people who attended East Coast schools *or* schools in provincial capitals.

To estimate the parameters α and λ, the procedure described in Chapter 3.2 was employed. The resulting estimated parameters were α = 2.8490 and λ = 0.1086. Table 3 shows the components of m^D, or the moments from the data that the model was calibrated with, and m^S, the model-predicted moments using these parameters:

Table 3: Population fractions found in the data and predicted by the model

Life path	m^D	m^S
f_1	0.8552	0.8641
f_4	0.0549	0.04826
f_5	0.0691	0.0690
f_6	0.01358	0.01581
f_8	0.00722	0.00399

The squared differences between the components of m^s and m^d sum to .00005948. After effectively calibrating the model, it is then possible to run policy experiments to simulate the effects of hypothetical policy changes.

5.1.2 Changes to the Hukou System

By varying the average cost of moving – which, due to the definition of the exponential distribution, is simply λ^{-1} – it is possible to simulate the effects of changes to the *hukou* system that would lower the mean of M by the same amount. For example, it is possible to calculate the fraction of the population choosing each life path given a hypothetical reduction of average moving costs by 10% or 25%:

Table 4: Population Fractions Following Two Possible Hukou Reforms

Life path	**Inititial m^s**	**m^s after 10% reduction in mean M**	**m^s after 25% reduction in mean M**
f_1	0.8641	0.8585	0.8466
f_4	0.04826	0.05318	0.06373
f_5	0.0690	0.06790	0.06555
f_6	0.01581	0.01736	0.02064
f_8	0.00399	0.00433	0.00504

It is also possible to calculate the overall fraction of the rural population that would choose to leave their homes (i.e. choose life paths f_4, f_6, or f_8, the sum of which I call the “rural migrant fraction”) given a particular average cost of moving. The results of these simulations are presented in Fig. 7:

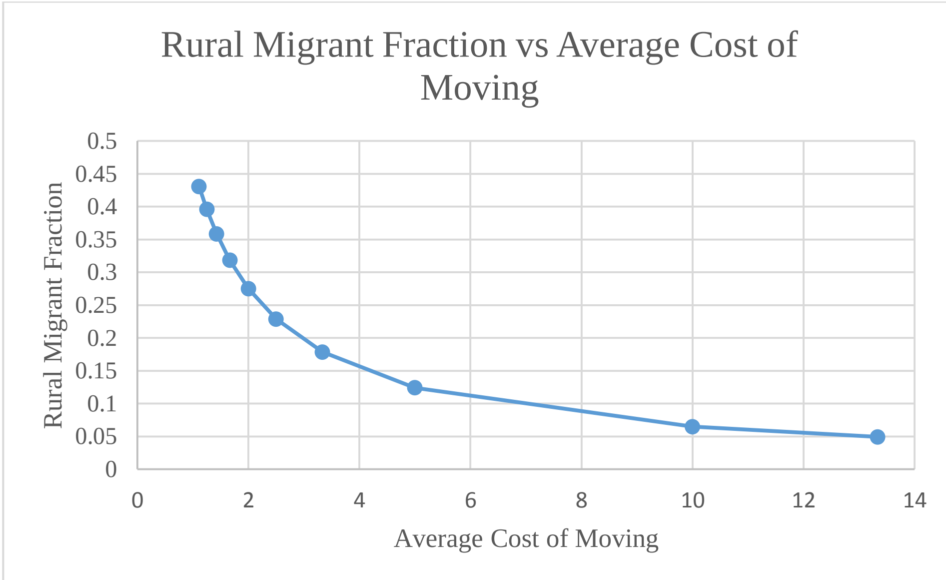

Figure 7: Changes to the Proportion of Rural Population Choosing to Migrate

Bear in mind that the current estimated average cost of moving is λ^{-1} = $(.108)^{-1}$ = 9.25 and the present rural migrant fraction is .0697. The relatively flat slope of the line presented in Figure 7 around the point (9.25, .0697) may at first glance suggest that small changes to the distribution of *M* will have a relatively small impact on the proportion of rural Chinese who choose to leave their homes, however to induce a migration such that that proportion increases by 30% to .0906, Chinese policymakers would have to lower the average cost of moving by 24.19% to 7.01, indicating an elasticity of migration greater than one along that range. As the hypothetical increases in the fraction of the rural population choosing to migrate increases, so too does the sensitivity of the rural population to changes in *M*. For example, to double the rural migrant fraction from its current value to .136, the average cost of moving would only have to fall by 51%. Whether such reductions are feasible is discussed in Chapter 5.2.

5.1.3 Changes to Admissions Policies

It is also possible simulate the effects of a changes to the structure of China's university admissions policies by varying the exogenous *gaokao* floor, i.e. θ^T. In the baseline model, θ^T was set to 1.5, which implies a situation in which entrance to top-tier universities is limited to rural-born Chinese in the top 1.4% of the ability distribution. Figure 8 below represents the way the college-bound rural population would respond to admissions policy reforms that effectively lower (or raise) this floor:

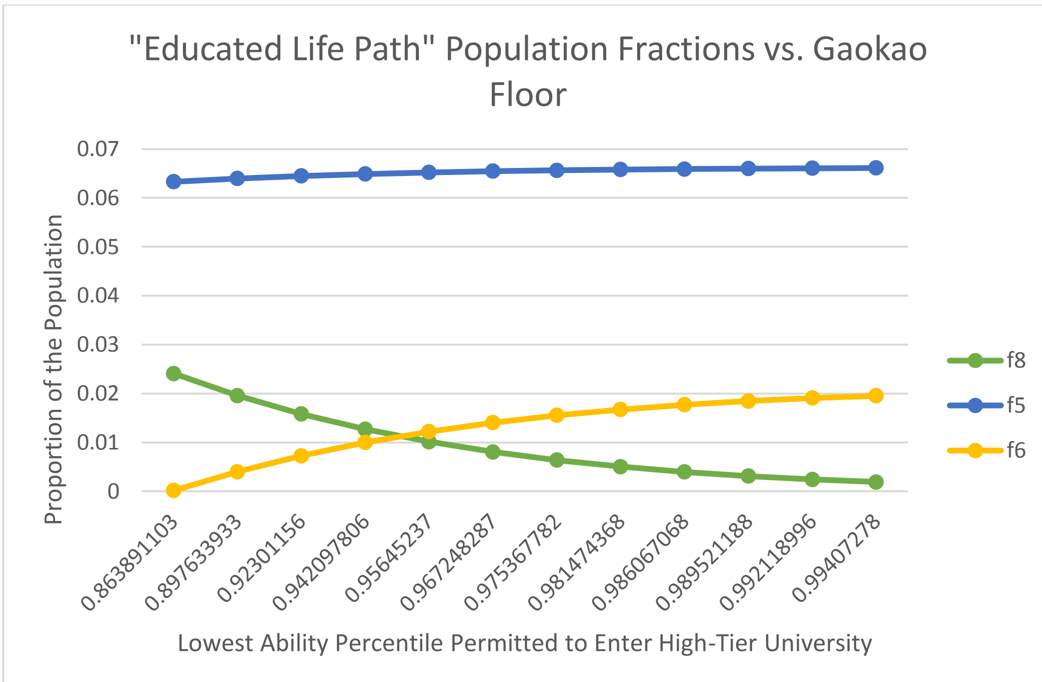

Figure 8: Changes to the Life Path Distribution in Response to Changes to the Gaokao Floor

As Figure 8 shows, the primary effect of lowering the *gaokao* floor is redistribution from f_6 to f_8, i.e. allowing progressively more rural Chinese with low M and high θ who would otherwise be forced to enroll in low-tier universities due to discriminatory treatment to matriculate into universities that better match their ability level. This shows the potential of a more equitable admissions system to reduce educational mismatch and thus improve the overall efficiency of the human capital accumulation process.

As the relatively flat line representing f_5 shows, however, this is not the only mechanism available to reduce mismatch. The primary beneficiaries from the redistribution resulting from a lower θ^T are those whose sole barrier to entry into a high-tier university is the very existence of that *gaokao* floor, that is to say those with low M and high θ. Lowering this barrier does this little to improve the lot of those with high M and high θ, who are forced into staying in the rural sector despite having ability levels that match or surpass those selecting life path f_8. Thus ameliorating the discriminatory way rural Chinese are treated in the country's university admissions process will not eliminate mismatch by itself. A significant portion of the high-ability rural-born population will only be able to enter a suitable university if their cost of moving is reduced (see Table 4 above for evidence of the effect of reducing average M on the size of f_5). Put another way, the most effective and equitable way to reduce educational mismatch is to accompany *gaokao* reform with *hukou* reform.

The sensitivity of the fraction of the population choosing f_8 with respect to the minimum ability percentile cutoff is quite high. For example, it is possible to increase the f_8 population by 60.91% by lowering θ^T such that the percentile floor drops from 98.6 to 97.5, indicating an elasticity of 54.87 along that range. Whether this reform is feasible is also discussed in the next section.

5.2 Discussion

The results of the policy experiments presented above should be interpreted as suggestions of the high degree of inefficiency and unfairness caused by these two policies and the incredible potential of reforms to these policies to improve the structure of the Chinese economy. That is to say that it would be imprudent to interpret the results as *precise* estimates of how policy changes would affect the rural life path distribution. In addition to the methodological issues already discussed, these experiments rely on the implicit assumption that all other relevant variables will remain constant despite to changes to the distribution of M or the *gaokao* floor θ^T. This, of course, is not very realistic. Among other changes, increased migration and high-tier university enrollment would likely lower the wage rate for urban skilled and unskilled labor, increase the wage rate for rural labor, increase urban housing prices, and, at least in the short-term, increase urban unemployment. All of these forces would act to retard further increases in migration and university enrollment.

It is also worth examining the feasibility of the reforms in question. What would it mean in practice to lower the average cost of moving by 24.19% (which the model predicts would increase the migrant fraction by 30%)? Remember that M is the sum of financial, psychological, and policy-induced costs, and the present average M is estimated to be equivalent to 9.25 times the average annual income of unskilled workers in the rural sector. Even with the fairly conservative assumption that the average financial and psychological costs of moving are each equivalent to the average peasant's annual income, that still implies that, at least for the average rural resident, 78.3% of moving costs are due to the *hukou* policy. That implies that increasing migration by 30% would require reducing policy-induced costs by 30.76%. While it may seem strange to conceive of the *hukou* policy in such numerical terms, many cities have point-based quota systems for determining which rural-to-urban migrants are eligible for local urban *hukou*. The model predicts that if, on the aggregate, these *hukou* "point floors" are effectively lowered to the extent that it becomes 30.76% easier to gain local urban *hukou*, then the rural-urban migration rate will increase by roughly the same amount. Of course, the *hukou* "naturalization" policies of the most prosperous cities, like Beijing and Shanghai, will have a disproportionate impact of these results.

In terms of reforms to the *gaokao* system, it was previously mentioned that lowering the ability percentile floor, below which no rural student can enroll in a high-tier university, by just

1.1% would result in an increase in the size of the rural student body at high-tier universities by 60.91%. Due to the "numerical" nature of the *gaokao* test, it is much easier to conceive of what this would look like in practice, with the added caveat that schools, unlike cities, have a finite number of slots available for incoming students. In the absence of investment in expanded facilities (at the expense of both scarce financial resources and school prestige, as the overall admissions rate would subsequently rise), an increase in the absolute size of the rural student body would have to be accompanied by an equal drop in the size of the urban student body. From an efficiency perspective, the urban students who would lose from such a change are those are "over-matched" with the quality of the school they are currently enrolled in, i.e. urbanites who only gained admission due to the discrimination against those from the countryside. This would improve the overall efficiency of the higher education system and, from an equity standpoint, make the system more meritocratic. At the same time, from a political economy standpoint, changes to both the *gaokao* and *hukou* system are very difficult to enact because the most powerful decision makers in China benefit personally from the privileges accorded to those with *hukou* registration in major cities and consider urban residents to be a major potential political threat (due to the more easily surmountable nature of the collective action problem in urban areas) who must be placated and coopted.

There are two main strengths to this method of research. First, unlike the vast majority of scholarly work on these subjects, these two policies are examined in the same context. Due to their deeply intertwined relationship, it is fairly unreasonable to analyze them in isolation. Next, the construction of a calibrated theoretical model allows the researcher to engage in comparative statics in order to make predictions about changes to the institutional environment and macroeconomic variables, something that is not possible in the absence of an explicit model. At the same time, as mentioned throughout this paper, constructing and calibrating this model required several semi-dubious assumptions. Some of these were only necessary due to imperfections in the data. Others, however, were required in order to make the model work at all, for example the assumption that there is no “reverse migration” from urban to rural areas. The net effect, is that this model is less of a crystal ball than it is a useful illustration of factors affecting labor mobility and human capital accumulation that explicitly incorporates two policies specific to the Chinese institutional environment.

Chapter 6

Conclusion

This paper evaluates the combined effect of the *hukou* and *gaokao* systems on the choices of China's rural population by constructing a calibrated, two period model of education and migration decisions and performing comparative statics to simulate possible reforms. The results of these policy experiments indicate that the household registration system and the policies that govern admission into China's universities place a significant burden on the country's rural populace and that even small changes to these policies could lead to large changes in the aggregate welfare of that group. From a normative perspective, this is just one more piece of evidence suggesting a moral imperative to promote greater equality of opportunity between China's rural and urban populations. From a pure efficiency standpoint, the results of this study suggest that *hukou* and *gaokao* act as two major distortions in the efficient allocation of labor and talent, with major implications for the overall productivity of the Chinese economy. Like all economic policy reforms, there would of course be tradeoffs involved in eliminating these burdensome policies. Overall, however, the net effect would likely be to improve both equity and efficiency, a rare remaining case of "low hanging fruit" that Chinese policymakers can and should pursue, assuming they can overcome the considerably large political barriers to doing so.

BIBLIOGRAPHY

Akgüç, M., Liu, X., & Tani, M. (2014). Expropriation with Hukou Change: Evidence from a Quasi-Natural Experiment. *IZA Discussion Papers*, No. 8289.

Au, C., & Henderson, J. V. (2006). How migration restrictions limit agglomeration and productivity in China. *Journal of Development Economics, 80*(2), 350-388.

Banerjee, A.V., and Moll, B. (2010). Why Does Misallocation Persist? *American Economic Journal: Macroeconomics*, *2*(1), 189-206.

Banerjee, A. V. & Duflo, E. (2005). Growth theory through the lens of development economics. in *Handbook of Economic Growth*, Aghion and Durlauf, ed. (1st ed., pp. 473-552) Elsevier.

Brandt, L., Hsieh, C., & Zhu, X. (2008). Growth and Structural Transformation of China. *China's Great Economic Transformation*, Brandt & Rawski, ed. Cambridge UP.

Brandt, L., Tombe, T., & Zhu, X. (2013). Factor Market Distortions Across Time, Space and Sectors in China. *Review of Economic Dynamics, 16*, 39-58.

Buera, F. J., Kaboski, J. P., & Shin, Y. (2011). Finance and development: A tale of two sectors. *The American Economic Review, 101*(5), 1964-2002.

Chan, K., Henderson, J., & Tsui, K. (2008). Spatial Dimensions of Chinese Economic Development. *China's Great Economic Transformation*, Brandt & Rawski, ed. Cambridge UP.

Cui, Y., Nahm, D., & Tani, M. (2013). Earnings Differentials and Returns to Education in China, 1995-2008.

Feenberg, D. R. & Poterba, J. M. (1993). Income Inequality and the Incomes of Very High-Income Taxpayers: Evidence from Tax Returns. *Tax Policy and the Economy*. Vol. 7. Poterba, J., ed., NBER, MIT Press.

Fu, Q., & Ren, Q. (2010). Educational Inequality under China's Rural-Urban Divide: The Hukou System and Return to Education. *Environment and Planning A, 42*(3), 592-610.

Harris, J. R., & Todaro, M. P. (1970). Migration, Unemployment and Development: A Two-Sector Analysis. *American Economic Review, 60*(1), 126-142.

Herd, R., Koen, V. & Reutersward, A. (2010). China's labour market in transition: Job creation, migration and regulation. *OECD Economics Working Papers*, No. 749.

Hsieh, C., Hurst, E., Jones, C. I., & Klenow, P. J. (2013). The Allocation of Talent and U.S. Economic Growth. *NBER Working Paper*, No. 18693.

Hsieh, C., & Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *Quarterly Journal of Economics, 124*(4), 1403 – 1448.

Ito, J. (2008). The removal of institutional impediments to migration and its impact on employment, production and income distribution in China. *Economic Change and Restructuring, 41*(3), 239-265.

Jeong, H., & Townsend, R. M. (2007). Sources of TFP growth: Occupational choice and financial deepening. *Economic Theory, 32*(1), 179-221.

Kamal, F., & Lovely, M. (2012). Labor Allocation in China: Implicit Taxation of the Heterogeneous Non-State Sector. *CESifo Economic Studies*, 1-28.

Knight, J., Song, L.,& Huaibin, J. (1999). Chinese rural migrants in urban enterprises: Three perspectives. *The Journal of Development Studies, 35*(3), 73-104.

Leal-Ordóñez, J. C. L. (2014). Tax collection, the informal sector, and productivity. *Review of*

Economic Dynamics, *17*(2), 262-286.

Liao, P., Wang, P., Wang, Y., & Yip, C. K. (2015). Educational Choice, Rural-Urban Migration, and Economic Development: The Role of *Zhaosheng* in China. NBER Working Paper.

Liu, Y., & Shen, J. (2014). Spatial patterns and determinants of skilled internal migration in China, 2000–2005. *Papers in Regional Science,93*(4), 749-771.

Meng, X., & Zhang, J.. (2001). The two-tier labor market in urban china occupational segregation and wage differentials between urban residents and rural migrants in Shanghai. *Journal of Comparative Economics,29*(3), 485-485.

Restuccia, D. & Rogerson, R. (2008a). Misallocation and Productivity. *Review of Economic Dynamics, 11*, 1-10.

Restuccia, D. & Rogerson, R. (2008b). Policy Distortions and Aggregate Productivity With Heterogeneous Establishments. *Review of Economic Dynamics*, *11*, 707-720.

The Economist (2015). A tightening grip: Rising Chinese wages will only strengthen Asia's hold on manufacturing. *The Economist.* Web.

Vollrath, D. (2009). How Important are Dual Economy Effects for Aggregate Productivity? *Journal of Development Economics, 88,* 325-334.

Wang, W. & Moffatt, P (2008). *Hukou* and Graduates' Job Search in China. *Asian Economic Journal, 22*(1), 1 – 23.

Yang, D. T., & Zhou, H. (1999). Rural-urban disparity and sectoral labour allocation in China. *The Journal of Development Studies, 35*(3), 105.

Zhong, H. (2011). Returns to higher education in China: What is the role of college quality? *China Economic Review, 22*(2), 260 - 275.

ACADEMIC VITA

EDUCATION:

The Pennsylvania State University **University Park, PA**
Schreyer Honors College *August 2012 – May 2016*
Bachelor of Arts in **Economics, International Politics, and Chinese**

East China Normal University **Shanghai, China**
CIEE Accelerated Chinese Language Program *June 2013 – August 2013*

- Completed the equivalent of one year of university-level language study in eight weeks

PROFESSIONAL EXPERIENCE:

The Brookings Institution **Washington, DC**
Research Intern, Project on International Order and Strategy *June 2015 – August 2015*

- Provided research support for Fellows' projects related to the breakdown of the post-Cold War international order, the "return of geopolitics," and U.S. grand strategy
- Wrote background memos, compiled source lists, and drafted outlines to assist with scholars' work on the European "pivot to Asia," the future of U.S. stability operations, Middle East geopolitics, and the effect of public opinion on U.S. foreign policy
- Helped organize and staff private Brookings events on topics such as Chinese naval strategy, the Greek financial crisis, and the international campaign against ISIS

Penn State Center for Global Business Studies **University Park, PA**
Research Assistant *January 2015 – May 2015*

- Conducted secondary research on the origins, implications, and determinants of success of China's efforts to internationalize the renminbi
- Synthesized research in order to draft CGBS reports and op-eds
- Led the development of a case study intended for an MBA-level international finance course designed to increase students' understanding of Chinese policymakers' thought processes vis-à-vis renminbi internationalization

U.S. Army War College **Carlisle, PA**
Intern, Peacekeeping and Stability Operations Institute (PKSOI) *May 2014 – August 2014*

- Worked with a Senior Foreign Service Officer to overhaul the curriculum of the War College's China elective, a course taken by senior leaders in the U.S. Armed Forces and exchange officers from other militaries
- Created the "course directive" for a similar course to be offered in a distance-education format

EXTRACURRICULAR ACTIVITIES:

Penn State Journal of International Affairs **University Park, PA**
Vice President, Co-Editor *April 2013 – May 2014; March 2015 – Present*

- Re-founded organization as vice president after it had lapsed into inactivity
- Active in staff recruitment, article solicitation, and formulating the club's long-term strategic plan
- Screens submissions, edits selected papers, and coordinates with authors

State of State **University Park, PA**
Projects Committee Member *September 2014 - Present*

- Creating tangible solutions to the issues raised at "State of State," an annual conference dedicated to improving Penn State
- Helping to launch "State Your State," a student-run blog intended to spark critical discussion about positive and negative aspects of the Penn State experience

HONORS AND AWARDS:

- Penn State Department of Economics Undergraduate Award
- Schreyer Honors College Delegate to 2014 Naval Academy Foreign Affairs Conference
- Phi Beta Kappa
- Evan Pugh Scholar (Top 0.5 Percent of Class)
- Paterno Liberal Arts Undergraduate Fellow

SKILLS:

- Mandarin Chinese (Limited Working Proficiency)
- Stata (Beginner)