

Left Behind by Design: School Access and Family Migration

Yang Tang

Nanyang Technological University

Jingwei Zhou

Nanyang Technological University

August 27, 2026

Abstract

Wherever public schools are financed locally and enrollment is tied to residence status, cities control migrant children's access to public education. China's hukou system makes this margin measurable. We study it in a spatial, dynastic equilibrium model in which families choose where to work, whether to bring their children, and how much to invest in them. Access enters as the institution works: a brought migrant child is enrolled by lottery, drawn only after the family has moved. The lottery matters: a certainty-equivalent spending discount cannot produce the excluded migrant children observed in the data, and it understates the family-arrangement response by a third. Equalizing access reunifies families even as migration rises, raises migrant children's human capital, and improves intergenerational mobility. At the pace of budget adjustment we estimate from three decades of city data, the reform redistributes across generations rather than across lineages. Natives of restrictive cities pay through a temporary budget dip, but every birth lineage gains in the long run; at the lower end of the estimated fiscal elasticity, the most protected lineage roughly breaks even. Restriction itself has few beneficiaries: opening the megacities delivers the largest national gain of any single-tier access reform, without hurting the megacity-born.

Keywords: hukou; migrant children; left-behind children; local school finance; intergenerational mobility; spatial equilibrium.

JEL codes: H75, I24, J13, J61, O15, R23.

1 Introduction

Urbanization in the developing world runs through families. When a worker moves to a city, the family faces a joint decision — whether the children come along, and where they will be schooled — and the destination faces one too: public schools are financed locally, so every migrant child admitted claims resources raised for someone else’s children. Wherever school finance is local and enrollment is tied to residence status, cities hold a policy margin over migrant children’s access to public education, and hundreds of millions of children worldwide grow up on one side of it or the other — brought to cities and schooled on unequal terms, or left behind. This paper asks what that margin does and who it serves: how a city’s restriction of migrant children’s enrollment shapes family arrangements, children’s human capital, intergenerational mobility, and welfare across birth lineages — and who gains, who loses, and who pays when the restriction is removed.

We study the question in China — not because the margin is Chinese, but because China is where it can be measured. The household-registration (*hukou*) system makes the restriction explicit and city-specific: education entitlements attach to a registered home locality, so migrant children’s access is an observable policy schedule rather than an implicit outcome of catchment rules — and the country’s census and household surveys record, at scale, the family arrangements, child outcomes, and city budgets the question turns on. The mechanisms our analysis isolates — enrollment risk, the family-separation margin, fiscal dilution, and skilled parents’ ability to self-insure privately — are not specific to the hukou institution that makes them measurable.

We begin with three facts that any analysis of this margin must confront. Access declines steeply with destination size: migrant children’s public-school enrollment falls from near parity in small cities to barely better than even odds in the megacities. Restrictions separate families, and they do so *selectively*: in the 2005 population census, the probability that a less-educated migrant mother has a child left behind rises by some twenty-one percentage points from the least to the most restrictive destinations, against roughly seven for more-educated mothers — and the gradient appears only for mothers whose children are of school age. And the relationship is not merely cross-sectional: family separation in the largest cities fell, relative to small cities, after a 2008 reform abolished the fees migrant children paid in urban public schools, and rose — by twelve percentage points — after a 2014 reform tightened access there by rule, the sign flipping with the direction of policy both times. Finally, school budgets adjust only partially to the student body: over three decades of city data, the elasticity of municipal education spending to enrollment is one half, so per-student resources dilute as migrant children claim seats.

We then build these three objects into a spatial, dynastic overlapping-generations model. Adults live a finite life cycle, work in one of several cities, and are Barro–Becker altruistic toward a single child whose adult human capital is produced during childhood from parental

time, private money, and public spending in the child's city of schooling; city wages move with size through an agglomeration elasticity, and the productivity of education rises with local average human capital through a Lucas/Bénabou externality. Access enters as the institution operates: a brought migrant child is *enrolled* with a city-specific probability — drawn after the family has chosen where to live and whether to bring the child — or *excluded* and schooled entirely at family expense. Around this enrollment lottery sit the other two instruments of the hukou menu: a settlement cost of keeping a child in the city, and education-scored residence conversion, by which skilled migrants acquire local status.

We calibrate the model to China's post-liberalization regime of the late 2000s and early 2010s. The economy has seven locations: the rural sector and the six official city-size classes into which China's roughly 290 prefecture-level cities fall, from small cities up to the megacities above ten million residents; the spatial objects — tier productivities, the birth distribution, bilateral migration costs — are built from the 2005 census and a prefecture-pair travel-time matrix. The access schedule is anchored to the published enrollment shares. The fiscal rule carries the budget elasticity we estimate in three decades of city budgets: education spending rises about half as much as enrollment, so half of each newly enrolled child's seat is financed by diluting everyone else's. And the urban skill bias — the extra return a unit of human capital earns in cities relative to the countryside, the force that makes migration skill-selective — is measured from the census returns gradient rather than fitted. The calibrated model reproduces a census fact it was never asked to match: the co-residence response to access is large at the bottom of the skill distribution and zero at the top, because skilled parents self-insure with private inputs — the same force that shows up in the data as the education gradient of family separation.

The lottery formulation itself is substantive, not cosmetic. We quantify what it buys by solving the entire model once more under the natural shortcut — a certainty-equivalent discount in which every brought migrant child is schooled with certainty and simply receives spending scaled by the enrollment probability (Section 5). The discount economy has no excluded children at all, where the data put 24% of brought migrant children outside public school with a quarter-standard-deviation skill gap. And exclusion as a discrete draw hits the family budget in the period it occurs: the family must replace the public input at the market's full cost. The same policy episode therefore moves family arrangements about one and a half times as much under the lottery as under the discount (+4.1 against +2.9 percentage points, a gap that widens as the bring choice sharpens), and the equal-access reform is worth a quarter more in entrant welfare (a 29% permanent-consumption equivalent against 23%). A contemporaneous parental valuation of child development adds roughly a further tenth to the response.

Equalizing access is a family-reunification and mobility reform, not a budget-shares adjustment. Along the perfect-foresight transition, families reunify and cities grow: co-residence of rural migrants rises 5 percentage points — most where restriction was harshest — while rural-urban migration itself rises, so removing enrollment risk pulls families into cities rather than

pushing them home. Migrant children's graduation human capital rises 28% in the long run after a transitory dip: megacity budgets take a per-student cut when every brought child claims a seat, but the cut fades — budgets adjust at the estimated, partial pace, and the reform's own in-migration rebuilds the urban tax bases — while the gains from newly enrolled and newly reunified children compound. The reform's redistribution is therefore intertemporal rather than dynastic — it moves welfare across generations within family lines, not from one line to another. The natives alive at the reform — adults living in their own birth tier when the policy arrives — pay, through the budget dip their children's schools absorb; every *birth lineage* — a family line indexed by its ancestral home city, valued from the start of adult life — gains in the long run, the rural majority most of all, and intergenerational mobility rises.

Who pays for the reform, and how much, is set by its financing, not by the access expansion itself. The reform bundles two decisions — who may enroll, and who pays — and we bracket the second across three fiscal rules that share an identical pre-reform economy (Section 5, Table 10). With rigid local budgets — spending does not respond to enrollment at all, so every newly enrolled child dilutes per-student resources one-for-one — the reform keeps most of its value (+15% of consumption) but the incumbents pay for it permanently: megacity spending per student never recovers, and the top lineages lose at entry. At the estimated pace of budget adjustment, the dip is transitory: megacity natives alive at the reform lose 2.4% of consumption, and every lineage gains in the long run — with the one caveat that at the lower end of the elasticity's estimated range the most protected lineage roughly breaks even instead of gaining. And with funding that follows each newly enrolled child — a tax increase of under two percentage points — there is no dilution at all: even megacity natives lose only a fraction of a percent at the reform date, and the reform's net welfare gain roughly doubles. The ranking is unambiguous: the access expansion creates the gain under any financing, and pairing it with funding that follows the child is what removes nearly all of its losers.

Comparing the instruments of the hukou menu yields two design lessons. First, under the estimated fiscal rule every instrument makes every birth lineage better off in the long run; this holds as long as the enrollment elasticity stays above roughly one third, the lower end of its estimated range. What separates the instruments is who pays along the way. Universal residence conversion, which bundles certain enrollment with the end of settlement costs, is the deepest reform — and the harshest on the generations alive at the reform, whose schools dilute within their lifetimes while the gains accrue to the unborn; a settlement-cost cut is the mildest and nearly transition-free. Second, coverage is decisive: the skill-targeted conversion that China actually operates — points-based systems and college admission — captures 4% of the universal policy's welfare gain at the top-decile threshold that points systems and college admission actually reach (42% at a median threshold) and *reverses* its mobility effect, entrenching skilled dynasties where universal conversion loosens transmission. Guaranteed access is worth most to families least able to self-insure, so selecting recipients on skill selects against value.

Finally, we ask who benefits from the restriction where it binds hardest: the megacities. The model has no political process, but it can measure what each birth lineage would gain or lose if megacity access alone changed. The answer is that the restriction’s supposed beneficiaries gain nothing from it. Once budgets adjust at the estimated pace, the spending that exclusion protects for megacity natives is offset by the tax base and human capital the turned-away migrants would have brought, so the megacity-born lineage is roughly indifferent to its own restriction. The nation, meanwhile, has its largest stake exactly there: opening the megacities alone is worth over half of the entire equal-access gain. So if something keeps megacity enrollment barriers in place, it is not the protection of incumbent children’s school budgets: by the model’s accounting, that protection is worth roughly nothing to the children it is meant to protect.

The paper speaks to four literatures. First, the quantitative-spatial work on China’s internal migration and the hukou system. [Tombe and Zhu \(2019\)](#) and [Ngai et al. \(2019\)](#) quantify migration barriers as wedges on workers — on mobility costs, land rights, and transfers. [Imbert et al. \(2023\)](#) move the family to the center, modeling the bring-versus-leave decision under hukou in a positive framework of consumption, housing, and remittances. And [Sieg et al. \(2020\)](#) bring in the fiscal side: they estimate a spatial overlapping-generations model in which hukou restricts migrants’ access to locally financed public goods — schooling included — with endogenous local tax and expenditure policy and a tax-financed residency reform, quantifying the fiscal externality migrants impose on city budgets. Second, the literature on migrant and left-behind children, which documents the stakes our model prices: outcome gaps by school access ([Lu and Zhou, 2013](#); [Chen et al., 2020](#)), the costs of parental absence ([Zhang et al., 2014](#); [Meng and Yamauchi, 2017](#); [Antman, 2013](#)), and quasi-experimental work in progress on the response of family migration to enrollment policy ([Guo et al., 2026](#)). Its structural counterpart is [Huang and Zhang \(2021\)](#), who build a spatial equilibrium model of the enrollment restriction itself — parents choosing a destination and whether to bring the child — in which relaxing the restriction benefits low-skill families most. Against these antecedents our contribution is specific. Access is a discrete *lottery* resolved after the location choice rather than a deterministic wedge or discount — the formulation that activates the family-arrangement margin. The school budget’s response to enrollment is *estimated* and carried into the model as its fiscal rule, so the fiscal channel — itself not new — operates at the pace the data display rather than an assumed one. Barro–Becker dynastic altruism carries intergenerational mobility as an equilibrium object through every experiment. The solution is also fully dynamic. Households save, and the enrollment draw hits the family budget in the period it occurs, so wealth and private inputs are the only buffers against exclusion. Every policy experiment is solved as a perfect-foresight transition, not only a comparison of steady states: who pays along the way, and for how long, is part of every answer. And the incidence of each tier’s access policy is computed lineage by lineage, delivering the constituency map of Section 7. Third, the school-finance literature descending from [Fernández and Rogerson \(1996, 1998\)](#) and [Epple and Romano \(1998\)](#), in which

locally financed schools plus residential choice generate sorting and inefficiency; [Eckert and Kleineberg \(2021\)](#) study education policy in spatial general equilibrium for the United States. Our setting adds the dimension those frameworks lack and developing economies have: a *status* margin, whereby access depends not on where a family lives but on where it is registered — so exclusion can persist at the destination and policy operates city by city. Fourth, the dynastic human-capital tradition ([Lee and Seshadri, 2019](#); [Anstreicher, 2024](#)) and the human-capital-externality and poverty-trap literature ([Lucas, 1988](#); [Bénabou, 1996](#); [Durlauf, 1996](#)), whose technology we embed in a multi-city economy with agglomeration ([de la Roca and Puga, 2017](#); [Combes and Gobillon, 2015](#)).

The remainder of the paper presents the motivating evidence (Section 2), the model (Section 3), the calibration (Section 4), the education access reform — the transition to equal access, with its incidence and decomposition (Section 5) — the alternative policy experiments across the hukou menu (Section 6), and the winners and losers from access restriction — the incidence of each tier’s access policy on local and national interests (Section 7) — before concluding. An online appendix collects the notation, model variants and robustness checks, the full empirical appendix, and computational details.

2 Motivating evidence

Three facts motivate the analysis: how school access varies across destinations, how families respond to it, and what happens to school budgets when migrant children enroll. The model that follows is built to reproduce them.

Fact 1: access declines with destination size. A migrant child’s probability of a public-school seat falls steeply across the city-size hierarchy. Published enrollment measures document the gradient: [Chen et al. \(2020\)](#) report that the share of migrant children in public schools reached 79.5% nationally by 2014 but only 41–70% at the elementary level in the largest destination cities (Guangzhou 41%, Shenzhen 46%, Beijing and Shanghai near 70%). Access is near parity in small cities and barely better than even odds in the megacities; this gradient is the policy variation behind everything that follows.

Fact 2: family arrangements respond to school access — selectively. If enrollment risk is a real cost of bringing a child, family arrangements should respond to it. Three exhibits establish that they do, each supplying what the others cannot. A census cross-section shows *who* responds: separation tracks destination restrictiveness, most steeply for the least-educated. Reform-based event studies show the response is *causal*: separation fell after a 2008 liberalization and rose after a 2014 tightening, city group by city group. And enrollment event studies

show the response runs *through the schools*: enrollment itself moves on education-specific policy dates, in both directions. We take the three in turn.

Separation tracks restrictiveness — steepest for the least-educated. Table 1 tabulates, from the 2005 one-percent population census, the probability that a migrant mother has at least one child left behind, by her destination’s city-size group and her own education. The grouping variable is size, but Fact 1 gives it a policy reading: the larger the destination, the lower a migrant child’s chance of a public-school seat, so moving up the size groups means moving up in restrictiveness (agricultural-hukou, inter-county migrant mothers aged 20–39 residing in cities or towns; $n = 18,214$ across 283 destination cities).

Table 1: Children left behind, by destination city-size group and mother’s education (2005 census)

Destination	P(any child left behind)		
	junior high or less	high school +	gap
Small prefectures (least restrictive)	0.41	0.45	−0.04
Other prefecture cities	0.64	0.56	+0.08
Major second-tier cities	0.62	0.49	+0.13
Beijing/Shanghai/Guangzhou/Shenzhen (most restrictive)	0.71	0.58	+0.13

Notes: Sample: women aged 20–39, currently married, agricultural hukou, hukou registered outside the county of residence, residing in a city or town, household head or spouse, at least one surviving child. A child is left behind if the mother’s surviving children exceed her co-resident children. Destination groups are the standard city-tier (size) classification; by the access gradient of Fact 1, restrictiveness rises with size. Census sampling weights. Cells are weighted shares. A probit version of equation (1) yields the same pattern: the average marginal cross-difference from the least to the most restrictive group is +11 to +15 percentage points. The interaction is stable to migration-distance controls (inter-province indicator, years since leaving), spouse presence and education, a fractional outcome, and replacing the restrictiveness groups with continuous log city size ($\hat{\beta}_2$ between +4.5 and +5.5 points throughout) — and it is concentrated where the restriction binds: for mothers aged 28–39, whose children are of school age, the interaction is +5.5 points (s.e. 2.0), against +2.7 (s.e. 3.4, n.s.) for mothers aged 20–27, whose children are mostly below school age.

Two features matter. Separation tracks restrictiveness: from the least to the most restrictive destinations, the left-behind probability of less-educated mothers rises by roughly 21 percentage points (7.2 per group step in equation (1)), against roughly 7 for the more-educated. And it does so *selectively*. To quantify the interaction, we estimate the linear probability model

$$LB_{imc} = \beta_1 R_c + \beta_2 (R_c \times LowEd_m) + \beta_3 LowEd_m + f(\text{age}_m) + \varepsilon_{imc}, \quad (1)$$

where LB_{imc} indicates that mother m in destination city c has at least one child left behind, $R_c \in \{1, \dots, 4\}$ indexes the destination’s tier group in the order of Table 1 — 1 small prefectures, 2 other prefecture cities, 3 major second-tier cities, 4 Beijing/Shanghai/Guangzhou/Shenzhen — a standard city-tier classification that we read as increasing in restrictiveness, since access falls with destination size (Fact 1), $LowEd_m$ indicates junior-high education or less, and f is a quadratic in the mother’s age; estimation uses census weights and standard errors are clustered

by destination city. The estimates are $\hat{\beta}_1 = +2.3$ percentage points per group step (s.e. 2.2) and $\hat{\beta}_2 = +4.9$ points (s.e. 1.8): educated mothers' family arrangements are almost insensitive to destination restrictiveness, while less-educated mothers' separation probability rises by 7.2 points per step — three times the gradient — and in the most open destinations the education gap even reverses. Restrictions, that is, separate precisely the families least able to substitute for the public school system on their own. These cross-sectional gradients are descriptive — migrants sort across destinations — but the pattern points to a family-arrangement margin that responds to enrollment risk, with a response that falls in parental skill: as if educated parents can substitute their own resources for the public school system, and less-educated parents cannot. The cross-section thus identifies the *shape* of the response — who it concentrates on; whether the margin responds *causally* to policy is a question it cannot answer. Two national reforms — one tightening access, one liberalizing it — can, and the next two exhibits use them.

Separation moves when policy moves. The 2014 National New Urbanization Plan tied residence policy to city size *by rule*: the largest cities were directed to strict, points-based admission of migrants, while smaller cities were directed to relax — an assignment of treatment that is mechanical in destination size. We estimate an event study across the 2005, 2010, and 2015 census waves on the sample of equation (1) extended to all destinations ($n = 54,996$):

$$\begin{aligned} \text{LB}_{imct} = & \sum_{g \in \{T1, T2\}} \sum_{s \in \{2010, 2015\}} \beta_{g,s} \mathbf{1}\{c \in g\} \mathbf{1}\{t = s\} + \gamma_g \mathbf{1}\{c \in g\} \\ & + \delta_t + f(\text{age}_m) + \text{LowEd}_m + \varepsilon_{imct}, \end{aligned} \quad (2)$$

where T1 is the first-tier destination group — Beijing/Shanghai/Guangzhou/Shenzhen, group 4 of equation (1), directed to strict points systems — T2 the major second-tier group (group 3, partially treated), and the omitted comparison group pools the other and small prefectures (groups 1–2, directed to relax); δ_t are wave effects with 2005 as base, and standard errors are clustered by destination prefecture (Appendix Table A8). The raw cells behind the estimates (Appendix Table A7) tell the story already. Between 2010 and 2015 — a period of massive secular reunification — the separation probability of less-educated migrant mothers fell by 53 percentage points in the smallest-prefecture destinations (0.68 to 0.14), by 33 points in the middle tiers, but by only 20 points in the four first-tier cities (0.84 to 0.65). The strictly treated first tier declines least by a wide margin, the most-relaxed smallest prefectures decline most, and the two middle groups sit within a point of each other. In the estimates, $\hat{\beta}_{T1,2010} = -5.2$ points (s.e. 2.3) and $\hat{\beta}_{T1,2015} = +11.8$ (s.e. 2.6; second tier +1.7 and +6.1; group-level randomization inference, a continuous-dose specification identified across all 287 prefectures, and bounds on the 2015 sample definition are in Appendix Table A9). The three census waves in fact contain *two* education-policy events, one in each direction, and separation moved with both. Between 2005 and 2010, the 2008 reform abolished the “borrowing fees” migrant children paid in urban public schools — a purely education-side liberalization, whose effect on enrollment appears in

the next exhibit. Access opened, and first-tier separation *fell* 5.2 points relative to the control group. Between 2010 and 2015, the 2014 plan tightened access in the largest cities. Access closed, and first-tier separation *rose* 11.8 points relative to the control group — a swing of 17 points from its path in the liberalization years. A confound built on smooth size-correlated trends in housing, labor demand, or congestion cannot produce this sign reversal on the policy dates: whatever drives the estimates must itself flip direction in 2008 and in 2014, which is what education policy did. A school-age placebo supports the liberalization response in point estimate and is mixed on the tightening, for an institutional reason. The liberalization’s separation response is twice as large among mothers aged 28–39, whose children are of school age (−6.1 against −2.7 for mothers of mostly preschool children; the interaction is −3.3, s.e. 2.3, so the two are not statistically distinguishable). The tightening loads on both groups — as it should. The points systems introduced in 2014 condition school admission on documentation and social-insurance histories that families must begin accumulating years before a child reaches school age. Anticipation therefore blurs the school-age margin for a rule change, but not for the spot abolition of a fee (Appendix Table A10). Family separation moved exactly where the policy moved, in the direction the policy moved it — in both directions — and ordered by the strictness of treatment.¹

Enrollment itself moves with policy. The children behind these family arrangements appear directly in the enrollment counts. In a panel of roughly 300 prefecture-level cities spanning three decades we estimate event studies of log primary enrollment on the same destination groups, around the two access-policy events that bracket the period — one in each direction:

$$\log S_{ct} = \sum_{g \in \{T1, T2\}} \sum_{s \neq s_0} \vartheta_{g,s} \mathbf{1}\{c \in g\} \mathbf{1}\{t = s\} + \alpha_c + \delta_t + \varphi \log P_{ct} + u_{ct}, \quad (3)$$

where S_{ct} is city c ’s primary enrollment in year t , P_{ct} its registered (hukou) population — absorbing the city’s own demographic trend — α_c, δ_t city and year fixed effects, and the base year s_0 is the year before each event (2007 for the fee abolition, 2013 for the tightening); standard errors cluster by city, and the full coefficient paths are in Appendix E.7. After the 2008 abolition of migrant “borrowing fees” — the education-specific liberalization that opened urban public schools to migrant children without charge — large-city primary enrollment rises steadily relative to small cities: +15 log points in the major second tier by 2013 (s.e. 2.9), on flat pre-2008 trends, with the same but noisier path in the first tier. After the 2014 tightening, the first tier’s relative rise — running at four to five log points per year through 2012 — stops abruptly, the path flat to negative thereafter. The enrollment margin thus responds in both

¹Two caveats. The number of migrant mothers observed in the first-tier cities also fell relative to other destinations — the extensive margin responded alongside the family-arrangement margin, so the composition of remaining migrants is not fixed. And the 2015 census dropped the hukou-nature question, so rural status in that wave is proxied by rural land-contract rights; the 2005–2010 comparison, on the uniform agricultural-hukou definition, is unaffected. Appendix E.7 details both.

directions, on education-specific policy dates; because the 2008 event is a school-fee reform and not a general hukou or labor-market reform, it ties the family-separation response above to *school access* specifically, rather than to the broader bundle of city policies that move with destination size. Enrollment barriers, in short, are not merely correlated with family separation across destinations; when they move, the schools fill or empty and families reunite or separate with them. Together the three exhibits divide the labor of Fact 2: the cross-section identifies who responds, the two-event study establishes that the response is causal — in both directions — and the school-system paths pin the operative margin to the schools themselves.

Fact 3: school budgets adjust only partially to enrollment. The incidence of access reform depends on what happens to per-student resources when migrant children claim seats. In the same city panel we estimate

$$\log E_{ct} = \beta \log S_{ct} + \gamma \log F_{ct} + \alpha_c + \delta_t + u_{ct}, \quad (4)$$

where E_{ct} is city c 's education spending in year t , S_{ct} its primary-plus-secondary enrollment, F_{ct} its total fiscal expenditure, and α_c, δ_t city and year fixed effects. Full adjustment of budgets to the student body is $\beta = 1$; $\beta < 1$ means spending *per student* falls as students arrive. Table 2 reports $\hat{\beta} = 0.51$ (s.e. 0.04): a city absorbing ten percent more students raises education spending by five percent, so per-student spending falls by roughly five percent. The stickiness is not a short-run artifact — the elasticity is 0.32 at five-year differences and 0.47 at ten-year differences (columns 2–3) — and it is primarily fiscal rather than physical: the same specification with *teachers* as the outcome yields an elasticity of 0.84 (column 4), so staffing follows students far more closely than money does. Because enrollment could itself respond to budgets, school quality, or local conditions, Appendix E.8 re-estimates the elasticity with two sources of predetermined variation: enrollment predicted from census cohorts already born (children aged 6–15 in year t were counted in the prior census), and migrant exposure interacted with the 2014 access tightening. Every policy-induced or demographic estimate lies *at or below* the panel elasticity: exogenous enrollment shocks meet budgets at least as sticky as the baseline assumes. Education budgets are sticky at the horizon migration operates on: enrolling migrant children dilutes the resources of the incumbent students who share the schools. The estimated elasticity is carried into the model as its fiscal rule, so the pace of budget adjustment in the policy experiments is the pace the data display.

Together the three facts fix what a model of this policy problem must contain: a size-declining access schedule anchored to published enrollment shares; a family-arrangement margin that responds to enrollment risk — causally, as the 2014 reform shows — with a response declining in parental skill; and a local school budget that dilutes as enrollment expands. The model of the next section builds in exactly these three objects.

Table 2: Adjustment of city education resources to enrollment

	(1) levels	(2) 5-year diff.	(3) 10-year diff.	(4) teachers
log K–12 enrollment	0.512 (0.044)	0.324 (0.034)	0.467 (0.040)	0.842 (0.020)
log fiscal expenditure	0.549 (0.033)	0.515 (0.038)	0.572 (0.034)	—
Outcome	log edu. spending	$\Delta \log$ spending	$\Delta \log$ spending	log teachers
Fixed effects	city, year	year	year	city, year
Observations	6,058	4,315	3,101	8,738

Notes: City panel, 299 prefecture-level cities, 1993–2022. Column (1) estimates equation (4) and controls additionally for log college enrollment (education budgets include higher education); columns (2)–(3) estimate it in long differences within city; column (4) replaces spending with the count of primary-plus-secondary full-time teachers. Standard errors clustered by city.

3 Model

3.1 Environment

An individual lives J periods and dies afterwards with certainty. The first J_c periods are childhood, during which no decisions are made; from age $J_c + 1$ through J the individual works and makes decisions. The economy consists of I locations, indexed $i \in \{1, \dots, I\}$, producing a single final good; we call them “cities” throughout, but in the quantitative implementation each urban location is a *city-size class* — an aggregate of the prefecture-level cities in a size tier — and every statement about a “city” in the results is a statement about a size class. An individual is described by a scalar stock of general human capital h ; the human-capital accumulation process is specified later.

At the beginning of age $J_c + 1$ the individual has a child.² The child is born in the parent’s *home city* i_b — the city where the parent himself was born — which need not coincide with the city where the parent currently lives and works. While of pre-working age, the child resides either with the parent or in the home city.

3.2 Human capital, preferences, and saving

All human-capital accumulation takes place in childhood: an individual who enters the labor force at age $J_c + 1$ with stock h retains h until death. The education technology is as follows: a child with current stock h_c residing in city $i_c \in \{i_b, i\}$ — either the home city i_b or the city i where the parent works — whose parent has stock h , accumulates next-period human capital

²Exactly one child per adult, born at labor-market entry: fertility is exogenous, population is stationary, and generations are evenly spaced. Fertility timing and heterogeneity — children arriving later in the parent’s career, or families with several children pooling enrollment risk — are margins the model abstracts from; both would interact with the bring decision, and we leave them to future work.

according to

$$h'_c = z \xi X^\omega, \quad \omega \in (0, 1), \quad (5)$$

where X is the schooling *input composite*,

$$X \equiv \left(m + \phi g(i_c, i_b) \right)^\phi \left(t \tilde{h}^{1-\delta} \bar{H}_{i_c}^\delta + \frac{(1-\phi)g(i_c, i_b)}{w_i} \right)^{1-\phi}. \quad (6)$$

Here m is the parent's money investment, $t \in [0, 1]$ the fraction of parental time devoted to the child, $g(i_c, i_b)$ the public education spending received by a child born in i_b and educated in i_c , and (z, ϕ, ω) technology parameters. The shock ξ is i.i.d. across children and periods, lognormal with $\ln \xi \sim N(0, \sigma^2)$, and identically distributed in every city.

The functional form embeds a transparent reading of public spending. A share ϕ of $g(i_c, i_b)$ is goods-like (materials, facilities) and complements the family's money input m ; the remaining share $1 - \phi$ hires instructional labor, which substitutes for parental instruction. Since w_i is the wage per efficiency unit of labor, $(1 - \phi)g/w_i$ is measured in *efficiency units of labor* — exactly the units of the parental term $t \tilde{h}^{1-\delta} \bar{H}_{i_c}^\delta$, a time fraction times efficiency per unit of time. The two instruction inputs therefore enter the composite in comparable units.³ Splitting the public input across the two composites in the same proportions $(\phi, 1 - \phi)$ as the Cobb–Douglas exponents makes public and private investments perfect substitutes, with g arriving pre-allocated in the parents' own cost-minimizing ratio — equivalently, as if schools allocated their budgets efficiently between teachers and materials. This assumption follows [Lee and Seshadri \(2019\)](#) and [Anstreicher \(2024\)](#). A further restriction is that money buys goods, not instruction; Appendix C relaxes it, letting households purchase instructional labor directly in every child state, and shows the paper's conclusions are unchanged. Finally, the outer exponent $\omega \in (0, 1)$ imposes diminishing returns on the total investment bundle — the childhood counterpart of Ben-Porath curvature, which in models with a separate young-adult investment stage appears there instead (Anstreicher, 2024). This curvature is what makes the child-investment problem concave, with interior money and time choices.⁴

The stock $\tilde{h}$ on which parental time operates *accumulates*: in the child's first period ($j = J_c + 1$) the child has no stock of its own yet, so $\tilde{h} = h$, the parent's human capital; in every later

³Converting at the wage of the parent's own labor market keeps the household-supplied and school-supplied efficiency units priced consistently within the household's problem. [Anstreicher \(2024\)](#) instead converts public time expenditure into hours using an economy-wide constant (mean parent human capital); a common national conversion factor is an alternative we leave to robustness.

⁴A deliberate asymmetry: parental hours are scaled by the effective stock $\tilde{h}^{1-\delta} \bar{H}_{i_c}^\delta$ — an individualized input that builds on what the child already knows — while public instruction enters unscaled, as a standardized classroom input. Beyond realism, this gives the technology a *public anchor* wherever public spending is received: an enrolled or home-city-schooled child with a low stock and no private inputs still obtains a positive, balanced input bundle. An *excluded* migrant child has no such anchor — its inputs are entirely family-supplied, so degenerate low-investment outcomes are possible in the excluded state; this is the model's counterpart of the migrant-school left tail, and it is why full exclusion is no longer ruled out by an Inada property alone (its cost is the finite enrolled–excluded value gap). The parent's own stock h still matters in every period through the budget constraint: time devoted to the child forgoes $(1 - \tau)w_i h^{\theta_i} t$ of earnings.

childhood period ($J_c + 1 < j \leq 2J_c$) the child learns from its own accumulated stock, $\tilde{h} = h_c$. Human capital thus builds on itself (dynamic self-productivity), so every period's investment and location matters, and the child's stock at graduation reflects its entire childhood, not only the last period. This stock is blended, Cobb–Douglas, with the *average human capital of the child's city*, $\bar{H}_{i_c}$, through the spillover weight $\delta \in [0, 1]$: the effective input is $\tilde{h}^{1-\delta} \bar{H}_{i_c}^\delta$. This is a Lucas/Bénabou human-capital externality (teachers, peers, and school quality reflect the local stock), and it makes high-human-capital cities educate everyone better — a human-capital counterpart to the spatial agglomeration in wages. $\delta = 0$ recovers pure self-productivity. The city average $\bar{H}_i$ is an equilibrium object, determined jointly with wages and public spending.

The received spending depends on whether the child is schooled at home, and — away from home — on a discrete *enrollment* outcome. A child educated in its home city ($i_c = i_b$) is always enrolled and receives the full per-child public spending of that city, $g(i_c, i_b) = g_{i_c}$. A child educated away from home ($i_c \neq i_b$) — i.e. after being brought to the parent's working city — secures a public-school seat with the city-specific probability $\rho_{i_c} \in [0, 1]$: *enrolled* ($a = 1$), it receives the full spending $g(i_c, i_b) = g_{i_c}$; *excluded* ($a = 0$), it receives no public input at all, $g(i_c, i_b) = 0$, and its schooling that period is financed entirely by the family — the public terms in (6) vanish and the input composite reduces to $X = (m)^\phi (t \tilde{h}^{1-\delta} \bar{H}_{i_c}^\delta)^{1-\phi}$. Enrollment is drawn i.i.d. each childhood period, *at the start of the period, after the previous period's location choices* — so parents decide whether to bring the child under enrollment risk, and the access probability ρ_i is the policy instrument of the paper: $\rho_i = 1$ is equal access and $\rho_i < 1$ the status-quo restriction. This makes exclusion a first-order event for the *parent's own budget*: replacing the public input requires private money at the market's full cost, so access risk matters directly for family decisions rather than only through the child's discounted future. Parental time, by contrast, requires *co-residence*: when the child lives apart from the parent ($i_c = i_b$ while $i \neq i_b$), the time input is unavailable ($t = 0$), and a left-behind child is educated with remittable money m and the (certain) home public input. Separation therefore carries a child-development cost in itself, in line with the evidence on parental absence (Zhang et al., 2014; Meng and Yamauchi, 2017). A migrant child thus faces two opposing forces — possibly higher public spending g_{i_c} in the destination city, but only probabilistic access to it; private money can partially substitute in the excluded state. Public spending, where received, enters both as a complement to private money ($m + \phi g$) and as a substitute for parental time (g/w).

Individuals obtain flow utility $u(c)$ from consumption of the final good. Households save in a risk-free asset paying an exogenous gross return $1 + r$.

During the child-rearing years a parent additionally derives flow utility from the child's current human-capital development, $\eta \log \bar{h}'_c$, where $\bar{h}'_c \equiv zX^\omega$ is (5) with the shock ξ stripped out. Since $\ln \xi$ has mean zero, this equals the expected log accrual $\eta \mathbb{E} \log h'_c$, so nothing hinges on the timing of ξ within the period. This *warm-glow* term is the standard paternalistic preference over child quality in family macroeconomics: the parent values the child's development *now*, as

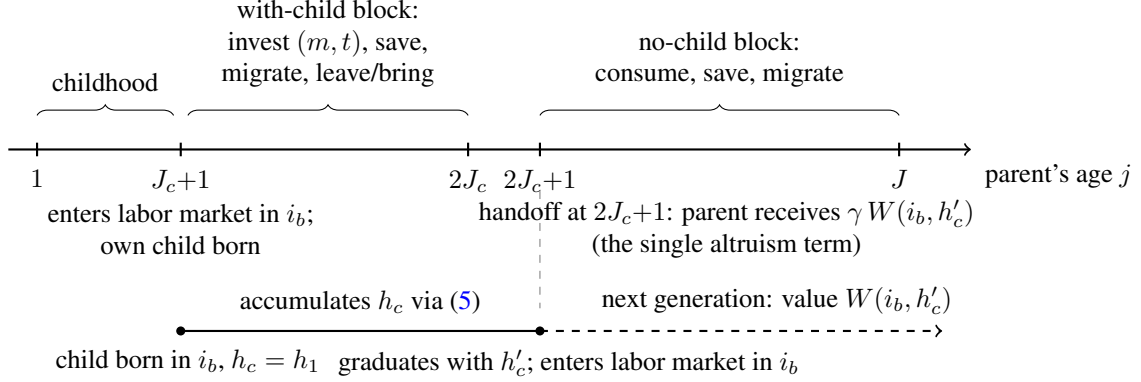


Figure 1: Life cycle and generational structure. The parent enters the labor market at age $J_c + 1$ with a newborn child; the with-child problem applies for ages $J_c + 1$ to $2J_c$; at $2J_c + 1$ the child returns to the lineage's birth city i_b and becomes the next generation's labor-market entrant, delivering the single altruism term $\gamma W(i_b, h'_c)$; thereafter the parent solves the no-child problem.

it happens. It is one of two motives the parent has for investing in the child; the other is dynastic altruism — the parent also values the child's entire discounted adult life, a term introduced with the household problem below. The two motives play distinct roles: the dynastic term carries the intergenerational welfare accounting, while the warm glow gives the parent a contemporaneous stake in the child's schooling. The glow is disciplined by the education-expenditure share in the calibration, and the paper's conclusions do not ride on it, as shown in Section 5.

3.3 Adult's Optimization Problem

Each working-age adult chooses, each period: (1) his own consumption c and saving b' , and — if the child is younger than working age ($j \leq 2J_c$) — the child's human-capital investment, consisting of a money input m and a time input t ; (2) the city $i' \in \{1, \dots, I\}$ in which to work next period, incurring a migration cost $d_{ii'}$ whenever $i' \neq i$; and (3) if the child lives apart from the chosen i' , whether to bring the child along. The migration cost $d_{ii'}$ and the settlement cost m^c introduced below are *utility* costs, subtracted directly from continuation values; only the education money input m enters the budget constraint. Figure 1 summarizes the life cycle and the generational structure; Appendix D collects the within-period sequence and the equilibrium feedback in a single diagram.

Adult without a dependent child. From age $2J_c + 1$ onward the adult solves a pure consumption, saving, and location problem: he chooses consumption c and saving $b' \geq 0$, given beginning-of-period financial wealth y (principal plus interest), so that next period's wealth is $y' = (1 + r)b'$. Labor earnings in city i are $(1 - \tau)w_i h^{\theta_i}$: w_i is the city's wage per efficiency unit, and $\theta_i \geq 1$ is a city-specific skill bias in efficiency units — $\theta_1 = 1$, so the rural sector is the reference — specified in Section 3.4, where it generates the model's migration selection. The

specification captures a robust feature of urban labor markets: the earnings advantage of cities is larger for more skilled workers (de la Roca and Puga, 2017). The next city is chosen subject to i.i.d. Gumbel preference shocks $\{\iota_{i'}\}$ with scale κ :

$$\begin{aligned} v_n(j, i, h, y) &= \max_{c, b' \geq 0} u(c) + \beta \mathbb{E}_\iota \left[\max_{i'} \{v_n(j+1, i', h, y') + \iota_{i'} - d_{ii'}\} \right] \\ \text{s.t. } \quad c + b' &= y + (1 - \tau) w_i h^{\theta_i}, \\ y' &= (1 + r) b'. \end{aligned}$$

Because the Gumbel shocks are independent, the continuation collapses to the log-sum-exp, and the corresponding city-choice policy is the logit share

$$\begin{aligned} \mathbb{E}_\iota \left[\max_{i'} \{\cdot\} \right] &= \kappa \log \sum_{i'} \exp \left(\frac{v_n(j+1, i', h, y') - d_{ii'}}{\kappa} \right), \\ \pi_n^{\text{move}}(i' \mid j, i, h, y') &= \frac{\exp[(v_n(j+1, i', h, y') - d_{ii'})/\kappa]}{\sum_{i''} \exp[(v_n(j+1, i'', h, y') - d_{ii'')/\kappa]}. \end{aligned}$$

Residence conversion. A migrant adult working away from home ($i \neq i_b$) obtains *urban residence status* with a per-period probability that rises with skill and depends on the work city,

$$p(i, h) = \min\{1, \bar{p}_i h^\psi\}, \quad \psi \geq 0,$$

where $\bar{p}_i$ is the city's base conversion level — the conversion rate at unit skill — and ψ the education-score slope. The h -dependence reflects the institution: urban residence in China is acquired through education-scored channels, points-based hukou systems and, above all, college admission. The cap binds the *product*, not the skill factor: $\bar{p}_i$ is a city-average rate over a mostly low-skill migrant pool, and high-skill migrants convert at rates approaching certainty in low-barrier cities; $\psi = 0$ recovers a skill-uniform profile. Conversion in the model is a *re-registration*: what it grants is a new home city, not a status that travels. How hard it is to obtain varies by city; a parent who later moves on faces the next city's enrollment lottery like any migrant, with the conversion city as the family's home base.

Conversion is drawn *each period* the adult works away from its home city, from the second working age on, at the rate $p(i, h)$ of the work city. With the draw, the family's home city becomes the work city: the child — brought to the parent, paying m^c once if it was left behind — is schooled as a native, with certain enrollment at the full local spending and no settlement cost, for as long as the family stays.

Conversion does not travel down the lineage in the benchmark: re-registration applies to the parent's remaining working life, and at graduation the child enters adult life at the ancestral birth city like every other child, so the entrant distribution $\{\pi_i^{\text{birth}}\}$ is invariant by construction. Section 6 evaluates the variant in which converted status passes to the child at the rate measured in the CFPS, making the birth distribution endogenous.

Adult with a dependent child. For ages $J_c + 1 \leq j \leq 2J_c$ the parent has a dependent child, and the state adds the home city i_b (initialized at the parent's birth city, re-registered on conversion), the child's current city $i_c \in \{i_b, i\}$, the child's human capital h_c , and — when the child is with the parent away from home — the period's enrollment status $a \in \{0, 1\}$ (for a child schooled in its home city, $a = 1$ always). Each period, events and decisions unfold in the following order.

- (0) **Status realizations.** From the second working age on, a family working away from its home city ($i \neq i_b$) converts with probability $p(i, h)$: its home city is re-registered to the work city (i_b becomes i), and the child, if at the old home city, is brought at the one-time cost m^c . If the family remains unconverted and the child resides in the parent's city away from home ($i_c = i \neq i_b$), the period's enrollment status is drawn: $a = 1$ with probability ρ_i , $a = 0$ otherwise.
- (1) **Consumption / saving / education.** Given the child's *current* city i_c and enrollment status a , the parent chooses consumption c , saving b' , and the child inputs (money m , time t ; the time input is available only under co-residence). The child's human capital accumulates to h'_c through the education technology (5) evaluated at i_c , with received public spending g_{i_c} if enrolled and 0 if excluded, and the spillover $\bar{H}_{i_c}$ of the child's current location.
- (2) **Migration.** Location preference shocks $\{\iota_{i'}\}$ are realized and the parent chooses next period's working city i' , paying migration cost $d_{ii'}$.
- (3) **Bring-child.** Given the chosen i' , a bring-child preference shock is realized and the parent chooses the child's next city $i'_c \in \{i_b, i'\}$: leave the child in the home city, or bring the child to the parent's city at cost m^c — *before* next period's enrollment is known.

The parent is *dynamically* altruistic (Barro–Becker): the child's entire adult life enters the parent's problem, discounted by γ , at the moment the child becomes an adult (the graduation term appears below, not as a per-period flow). Collecting step (1) and the continuation over steps (2)–(3) — with the enrollment status a in the state, degenerate at $a = 1$ for a child schooled in its home city —

$$\begin{aligned}
v_c(j, i, i_b, h, y, i_c, a, h_c) &= \max_{c, b' \geq 0, m, t} u(c) + \eta \log \bar{h}'_c \\
&\quad + \beta \mathbb{E}_{\iota, \xi} \left[\max_{i'} \left\{ \mathcal{A}(j+1, i', i_b, h, y', h'_c) + \iota_{i'} - d_{ii'} \right\} \right] \\
\text{s.t. } c + b' + m &= y + (1 - \tau) w_i h^{\theta_i} (1 - t), \\
y' &= (1 + r) b', \\
h'_c &\text{ given by the education technology (5), } \bar{h}'_c = z X^\omega, \\
\tilde{h} &= \begin{cases} h & j = J_c + 1 \\ h_c & j > J_c + 1 \end{cases}
\end{aligned}$$

where $\max_{i'}\{\cdot\}$ is the step (2) migration choice and $\mathcal{A}$ is the *arrival value* in the chosen city, defined below: the bring-child choice integrated out (the object V_c , distinct from the period value v_c), mixed over the conversion draw. The child affects the parent's payoff only through h'_c , which drives the child's value at graduation.

The bring-child decision. The continuation in the Bellman equation nests the two end-of-period choices — migration, then bring-child; we unfold them from the inside out, starting with the innermost, step (3). Consider an age- j parent working in city i' with wealth y and a child of stock h_c . Leaving the child home delivers a certain outcome, while bringing it exposes the family to next period's enrollment draw, which is *not yet realized* at the bring node — so the bring option is valued in expectation:

$$\begin{aligned} v_c^{\text{leave}}(j, i', i_b, h, y, h_c) &\equiv v_c(j, i', i_b, h, y, i_c = i_b, 1, h_c), \\ v_c^{\text{bring}}(j, i', i_b, h, y, h_c) &\equiv \rho_{i'} v_c(j, i', i_b, h, y, i_c = i', 1, h_c) \\ &\quad + (1 - \rho_{i'}) v_c(j, i', i_b, h, y, i_c = i', 0, h_c) - m^c \mathbf{1}\{i' \neq i_b\}, \end{aligned}$$

the first the certain leave option, the second the bring option with the enrollment risk borne. The marginal value of access at this node is the discrete enrolled–excluded gap, $\partial v_c^{\text{bring}} / \partial \rho_{i'} = v_c(\cdot, 1, \cdot) - v_c(\cdot, 0, \cdot)$ — an inframarginal object that includes the full budget cost of self-financing the public input, which is what gives access policy first-order purchase on the family-arrangement margin. The parent draws i.i.d. Gumbel bring-child shocks $(\zeta_{\text{leave}}, \zeta_{\text{bring}})$ with scale κ_c and takes the larger of $v_c^{\text{leave}} + \zeta_{\text{leave}}$ and $v_c^{\text{bring}} + \zeta_{\text{bring}}$. Integrating out the shocks gives the smoothed value and the bring-child probability

$$\begin{aligned} V_c(j, i', i_b, h, y, h_c) &= \kappa_c \log \left[\exp\left(\frac{v_c^{\text{leave}}}{\kappa_c}\right) + \exp\left(\frac{v_c^{\text{bring}}}{\kappa_c}\right) \right], \\ \pi_c^{\text{bring}}(i', i_b, \cdot) &= \frac{\exp(v_c^{\text{bring}} / \kappa_c)}{\exp(v_c^{\text{leave}} / \kappa_c) + \exp(v_c^{\text{bring}} / \kappa_c)}. \end{aligned}$$

When the parent is in the home city ($i' = i_b$) the two options coincide and the child is necessarily at home ($i_c = i_b$), so no settlement cost is incurred. In the Bellman equation above, the continuation is this object evaluated one period ahead, at age $j + 1$ with next period's wealth and child stock.

The converted family. Conversion introduces no separate value function: a converted family is an ordinary family whose home city has been re-registered to its work city. At home the leave and bring options coincide — the child lives with the parent and is schooled with certain enrollment at the full local spending, with no lottery and no settlement cost — so the family's problem is v_c evaluated at $i_b = i$. If the parent later moves on, the family is again a migrant family, now with the conversion city as its home: the child can be left there and schooled with certainty, or brought to the new city against its enrollment lottery, and a new conversion draw applies there. Note the difference from a favorable enrollment draw: an enrolled unconverted

family redraws the lottery next period and pays the settlement cost at future bring nodes, while re-registration removes both for as long as the family stays and upgrades the family's fallback from the rural home to an urban one if it moves — that is conversion's option value. The graduation handoff below is unaffected: the lineage attribute is the ancestral birth city, carried unchanged through re-registration.

The migration decision. Moving one layer out: at the start of the next period in a city i' away from home, the family faces the conversion draw of step (0), so the value of arriving is the mixture

$$\mathcal{A}(j+1, i', i_b, h, y', h'_c) = p(i', h) V_c(\cdot; i_b=i') + (1 - p(i', h)) V_c(\cdot; i_b),$$

where both V_c terms are evaluated at the same arrival state $(j+1, i', h, y', h'_c)$: the first with the home city re-registered to i' (where the bring options coincide), the second unconverted. Here $p \equiv 0$ at $i' = i_b$ and at the entry age $J_c + 1$ (conversion begins at the second working age), so at home $\mathcal{A} = V_c$. The step-(2) next-city choice is then a Gumbel discrete choice over i' with logit share

$$\pi^{\text{move}}(i' \mid j, i, i_b, h, y', h'_c) = \frac{\exp[(\mathcal{A}(j+1, i', i_b, h, y', h'_c) - d_{ii'})/\kappa]}{\sum_{i''} \exp[(\mathcal{A}(j+1, i'', i_b, h, y', h'_c) - d_{ii'')/\kappa]}.$$

The two schedules π^{move} (over cities) and π_c^{bring} (over child location) are precisely the policy objects that the stationary-distribution transition iterates on; the no-child value v_n uses the analogous city-choice share with v_n in place of V_c .

Dynastic altruism, entry, and the generational handoff. *Birth city is a fixed lineage attribute.* A child is born in its parent's ancestral home city i_b , and when it grows up its own home city is again i_b — residence conversion benefits the family within its lifetime but does not travel down the lineage. The cross-city distribution of labor-market entrants is therefore exogenous and constant, $\{\pi_i^{\text{birth}}\}$.

Entry. An adult enters the labor market at age $J_c + 1$ in its birth city i — with a newborn child of its own — and freely chooses a work city i' , migrating from home at cost $d_{ii'}$, together with whether to bring the child. Let $W(i, h)$ denote the ex-ante lifetime (dynastic) value of a labor-market entrant with birth city i and human capital h :

$$W(i, h) = \kappa \log \sum_{i'} \exp\left(\frac{V_c(J_c+1, i', i, h, 0, h_1) - d_{ii'}}{\kappa}\right),$$

where $V_c(J_c + 1, i', i, h, 0, h_1)$ is exactly the bring-child-smoothed value defined above, evaluated at entry: age $J_c + 1$, work city i' , birth city i , zero wealth, and a newborn child with the common initial stock h_1 — the leave/bring choice being between leaving the newborn at home i and bringing it to i' .

Handoff. For interior ages $J_c + 1 \leq j < 2J_c$ the child remains dependent and the bring-child node returns the next-period parent-with-child values. At the end of the parent's last with-child period ($j = 2J_c$) the child graduates — carrying the stock h'_c produced by that final period's education technology (5) — and *returns to its birth city* i_b to begin adult life there; there is no bring choice at graduation. The parent's age- $(2J_c + 1)$ continuation is

$$v_n(2J_c + 1, i', h, y') + \gamma W(i_b, h'_c),$$

the parent's own no-child value (having migrated to i') plus the dynastic value of the child. *This is the only point at which altruism enters the parent's lifetime utility:* the single term $\gamma W(i_b, h'_c)$, received at age $2J_c + 1$, capitalizes the child's entire adult life — and, recursively, that of all further descendants. During childhood the child affects the parent's problem only through the state h_c and the budget; from age $2J_c + 1$ on the parent solves the no-child problem v_n , which contains no altruism term.

3.4 City labor markets: agglomeration and congestion

Each city produces the final good — the numeraire, with price normalized to one — with a technology that is linear in efficiency labor,

$$Y_i = A_i N_i, \quad w_i = A_i,$$

where N_i is aggregate efficiency labor employed in city i : competition drives profits to zero, so the wage per efficiency unit equals productivity and the wage bill exhausts output. Agglomeration enters through the productivity level. Each firm takes A_i as given, but at the city level it moves with size,

$$A_i = \bar{A}_i N_i^\lambda.$$

The single elasticity λ governs the net effect of city size on the wage: $\lambda > 0$ is net agglomeration (larger cities pay more), $\lambda < 0$ net congestion (larger cities pay less). Aggregate efficiency labor in city i is

$$N_i = \int \mathbf{1}\{\nu \text{ works in } i\} h_\nu^{\theta_i} d\mu(\nu),$$

where the indicator is determined by the Gumbel location choice, so that city sizes — and hence wages — respond smoothly to the wage vector. Together with the migration costs $d_{ii'}$, this externality provides the dispersion (or concentration) force that pins down the distribution of population across cities in equilibrium.

3.5 Government

Public education is tax-financed and locally provided. Each city’s education budget responds to its student body according to the allocation rule

$$E_i = s_E w_i N_i \left(\frac{K_i}{\bar{K}_i} \right)^{\beta_g}. \quad (7)$$

Here the enrolled headcount $K_i = \Xi_i^{\text{nat}} + \rho_i \Xi_i^{\text{mig}}$ counts all native children (Ξ_i^{nat} , with $i_c = i_b$) plus the fraction ρ_i of migrant children who draw a seat; the budget is divided equally among them, so per-enrolled-child spending is $g_i = E_i/K_i$; $\bar{K}_i$ is a fixed reference enrollment; and β_g is the elasticity of the education budget with respect to enrollment, the object measured in Fact 3 of Section 2. Every enrolled child receives the same g_i ; excluded migrant children receive nothing. The rule nests the polar fiscal regimes — fixed budgets at $\beta_g = 0$, funding that follows each child at $\beta_g = 1$ — and Section 4 interprets it, sets $\bar{K}_i$ and β_g , and states what is estimated and what is imposed.

The national education budget must balance. Because education finance in China runs heavily through intergovernmental transfers, we treat the education system as nationally financed and locally provided: households everywhere pay the common labor-income tax rate τ that balances the national education budget,

$$\tau = \frac{\sum_i E_i}{\sum_i w_i N_i}, \quad (8)$$

so the tax rate is an equilibrium object. The budget share s_E in (7) is a parameter of the allocation rule — the share of the wage bill that anchors education budgets, with its value set in the calibration. The allocation rule spreads the national budget across cities; (8) governs the size of the budget itself: any fiscal increment a reform generates is paid by every worker through the tax, not received from outside the model. At the benchmark $K_i = \bar{K}_i$, so $\tau = s_E$ and the calibrated equilibrium is invariant to β_g , which governs only how budgets respond to reform. The balanced-budget requirement prices the resource cost of education expansion through household budgets; what it abstracts from is substitution against other local public services, which are outside the model. Appendix I describes how the policy experiments are computed under it. The schedule $\{g_i\}$ is an *equilibrium object*, determined jointly with wages, city sizes, and the distribution of children: raising migrant access ρ_i enrolls more migrant children at full per-child cost and dilutes per-child spending — at the estimated, partial rate — where the migrant-child load is large relative to the native one. The dilution–access trade-off at the heart of the policy analysis thus operates on enrollment headcounts, at the pace of adjustment the data display.

3.6 Stationary equilibrium

Given the policy $(s_E, \{\rho_i\})$, a stationary competitive equilibrium consists of a city-wage vector $\{w_i\}$, per-enrolled-child education spending $\{g_i\}$, a tax rate τ , city-average human capital $\{\bar{H}_i\}$,

value functions $\{v_n, v_c\}$ and associated policies, and a stationary distribution μ over adult states such that:

1. given $\{w_i\}$, $\{g_i\}$, $\{\bar{H}_i\}$, the return r , the tax τ , and the access schedule $\{\rho_i\}$, the value functions and policies solve the adults' dynamic-programming problems, including the expectations over the human-capital shock ξ and the Gumbel location and bring-child shocks;
2. μ is invariant under the transition induced by these policies — location, bring-child, saving, and the arrival of the next generation (birth at the parent's age $J_c + 1$, graduation to a new adult at $2J_c + 1$);
3. city labor markets are consistent, $w_i = \bar{A}_i N_i^\lambda$ with N_i aggregated from μ as above;
4. education budgets are consistent: g_i satisfies the allocation rule (7) with the child masses $\Xi_i^{\text{nat}}, \Xi_i^{\text{mig}}$ aggregated from μ , and the tax rate balances the national education budget (8);
5. the spillover is consistent: $\bar{H}_i$ equals the average adult human capital in city i under μ ; and

Table A1 in Appendix A collects the model's notation for reference.

4 Calibration

We calibrate the model to China's post-liberalization regime of the late 2000s and early 2010s, anchoring the spatial objects in 2005 (Appendix F discusses the vintages). Two kinds of data are used throughout: the 2005 one-percent population mini-census and prefecture-level yearbook aggregates, which provide city wages, registered population, government education spending, and migration flows; and the China Family Panel Studies (CFPS), a nationally representative household panel conducted by Peking University, which provides the household-level moments (education spending, income dispersion, co-residence of migrant children, and parent-child outcomes). We set the number of locations to $I = 7$: tier 1 is the rural sector, and tiers 2–7 group China's prefecture-level cities into the six official urban size classes by census urban population — small (<0.5 million), medium (0.5–1 million), Type-II large (1–3 million), Type-I large (3–5 million), super-large (5–10 million), and megacities (>10 million: Beijing, Tianjin, Shanghai, Guangzhou, Shenzhen, and Chongqing). A tier in the model is thus a class of similar-sized cities, ordered by size, with tier-level productivities, spending, and migration costs constructed as population-weighted aggregates of the underlying prefectures. Calibration then proceeds in two steps: *predetermined* parameters, taken from the literature or measured directly from the data without solving the model, and *jointly calibrated* parameters chosen so that the stationary equilibrium matches targeted moments.

4.1 Predetermined parameters

Life cycle and preferences. A model period is seven years. Individuals live $J = 8$ periods: a childhood of $J_c = 3$ periods (birth to age 21), during which no decisions are made, followed by five working periods (ages 21–56). The child is born at the start of the parent’s first working period, so generations are 21 years apart. Utility is logarithmic, $u(c) = \log c$: with utility-additive location-choice shocks and migration costs, log consumption makes the choice structure scale-free — proportional wage differences are worth the same to every skill level — which is the natural benchmark for migration selection (and the workhorse specification of the quantitative spatial literature). Curvature above one would make utility bounded and mechanically shut down skilled migration against fixed moving costs. The per-period discount factor $\beta = 0.70$ and risk-free return $r = 0.15$ correspond to roughly 0.95 and 2% at annual rates.

Parameters taken from the literature. The agglomeration elasticity is $\lambda = 0.06$, in the range surveyed by [Combes and Gobillon \(2015\)](#); it applies to the urban tiers only, with the rural wage fixed at $\bar{A}_1$. These are wage elasticities gross of housing and commuting costs, so the benchmark treats urban congestion as zero — an upward bias on the gains from concentration — and Appendix H therefore also reports the reform at $\lambda = 0$ and at a net-congestion value $\lambda = -0.05$ (Tables A16 and A17). The education curvature $\omega = 0.60$ sits in the middle of the range of Ben-Porath curvature estimates, roughly 0.5–0.9 ([Browning et al., 1999](#)); [Anstreicher \(2024\)](#) estimates 0.38 in a spatial model in which this curvature belongs to a separate young-adult investment stage. The money share $\phi = 0.40$ and the spillover weight $\delta = 0.30$ are set a priori, with δ motivated by the human-capital-externality literature ([Lucas, 1988](#); [Bénabou, 1996](#)). The altruism weight $\gamma = 0.50$ is fixed externally (it is collinear with the education scale z); its level lies in the standard Barro–Becker range of per-generation altruism weights.

The city-choice Gumbel scale is normalized to $\kappa = 2$; this only fixes the utility units of the location shocks. The bring-child scale κ_c cannot be normalized in the same way: with units already fixed, it is a substantive parameter — the dispersion of family-specific settlement costs around m^c — and it governs how sharply the family-arrangement margin responds to changes in the value of bringing the child. No targeted moment pins it down separately from m^c , so we set $\kappa_c = 0.5$; halving or doubling it, with m^c re-fit to the co-residence target in each case, leaves the paper’s conclusions in place.⁵

The education-shock dispersion. The dispersion of the idiosyncratic shock to child human capital, $\sigma = 0.30$, is disciplined by intergenerational persistence rather than by the dispersion target — the standard practice in dynastic human-capital models, where the size of the luck

⁵Across that range the equal-access reform’s welfare gain stays within $\pm 25\%$ of its central value, its reunification response moves by well under a percentage point, and its mobility gain — the fall in the IGE — stays between 0.07 and 0.09. The κ_c rows of Table 11 show the lottery-versus-discount comparison across the same range.

component is identified from how strongly child outcomes track parents (Restuccia and Urrutia, 2004; Lee and Seshadri, 2019). The dispersion target cannot identify it here: the structural sources — the enrollment lottery, family separation, and the skill-biased premium — account for observed income dispersion on their own, so an unrestricted σ is driven to zero. At $\sigma = 0.30$ the model’s intergenerational persistence lands within the empirical income-IGE range of 0.4–0.6 (Fan et al., 2021), at the cost of a modest overshoot of the dispersion target (Section 4.3).

The status-quo access schedule. The schedule $\{\rho_i\}$ is *anchored to published enrollment shares, and declines with city size*: under the enrollment-lottery formulation of Section 3, ρ_i is literally the probability that a migrant child secures a public-school seat in tier i , so its empirical counterpart is the public-school enrollment share e_i among migrant children. Not every tier value is observed: the megacity value is a migrant-weighted average of four published city shares, the rural value is one by construction, and the five intermediate tiers are assigned by interpolation between the observed city shares and the published national average, the whole schedule then rescaled slightly to match the aggregate (Appendix F details the construction; Table A14 shows the conclusions hold across two alternative schedules).⁶ The enrollment share falls steeply with destination size. Chen et al. (2020) report that it reached 79.5% nationally by 2014, after a decade of liberalization from below one half in the early 2000s — but only 41–70% at the elementary level in the largest destination cities (Guangzhou 41%, Shenzhen 46%, Beijing and Shanghai near 70%). The tier-level enrollment shares give the benchmark schedule reported in Table 3 — near-certain enrollment in the small tiers and barely-better-than-even odds in the megacities. Under the model’s own distribution of migrant children across tiers the schedule’s weighted mean is 0.70 (Appendix F), below the published national share of 0.795 because migrant children concentrate in the restrictive tiers; the schedule matches the post-liberalization regime that also generates our CFPS targets. Under the stricter early-2000s access regime the national share was nearer 0.5; Appendix F re-solves the model and the reform under both that schedule and a post-2014 one, and the paper’s conclusions hold across the range. The model’s implied outcome gaps for migrant children are in line with the test-score evidence in Lu and Zhou (2013).

Skill bias and residence conversion. Two objects introduced in Section 3 complete the pre-determined set. The urban skill bias — the exponent θ_i in earnings $(1 - \tau) w_i h^{\theta_i}$, which makes skill better paid in cities — takes a common value θ_u in every urban tier ($\theta_1 = 1$), anchored

⁶Enrolled migrant children may receive slightly less funding than natives where “out-of-district” fees and informal costs apply (a funding share $\varphi^{\text{pub}} \approx 0.9$); we abstract from this second-order wedge and give enrolled children full g_i . Excluded children attend migrant-run schools whose public funding is close to zero. Modeling the enrollment margin as an explicit lottery follows the institutional reality: enrollment is decided by the destination system after the family has moved. The certainty-equivalent alternative — discounting spending to ρg — is solved and compared in Section 5 (Table 11): it removes the excluded state entirely and mutes the family-arrangement response by a third.

to the urban-to-rural gradient of returns to schooling in the 2005 census. Among agricultural-hukou workers — the composition-held comparison the model’s location premium calls for, since it prices the same population in different work locations — the Mincerian return to a year of schooling is 6.3% in cities against 4.7% in the countryside, a ratio of 1.34; we set $\theta_u = 1.35$.⁷ Because θ_u is measured rather than fitted to any migration moment, migration selection remains available as a genuine over-identification check (Section 4.3). The residence-conversion levels $\bar{p}_i$ are the measured hukou-conversion registration shares by tier — zero in the rural tier, declining from 6% to 1% per period across the urban tiers — and the education-score slope is set to $\psi = 2$. ψ is set rather than calibrated because it has no leverage on any candidate moment: the top-end selection share is flat in ψ across $[0.5, 8]$ — parents facing enrollment risk self-insure through private inputs, so a prize denominated in school access cannot sort adults by skill — and conversion is accordingly retained as a policy margin (the third instrument of the hukou menu, alongside ρ_i and m^c), not as the model’s selection device.

The fiscal rule. The budget-adjustment elasticity is set to its estimated value, $\beta_g = 0.51$ (Fact 3 of Section 2): city education budgets absorb roughly half of an enrollment change, at short and ten-year horizons alike (Table 2); instrumented estimates lie at or below this value (Appendix E.8). The reading is budgetary incrementalism — budgets respond at rate β_g to deviations from the student body they have historically served — so $\bar{K}_i$ plays the role of the city fixed effect in the estimating regression, and the calibration sets it to benchmark enrollment; the rule then passes through the observed budget–enrollment point. Because Fact 3 controls for total fiscal expenditure, β_g measures how the education line-item moves *within* a city’s budget; the budget’s total size is governed separately, by the balanced-budget condition (8). One restriction is imposed rather than estimated: the budget carries a unit elasticity with respect to the local wage bill $w_i N_i$; Table A17 re-runs the reform with the estimated exponent (0.55) instead. Because $\bar{K}_i$ is set to benchmark enrollment, the calibrated equilibrium is invariant to β_g ; the elasticity matters only for the policy experiments, which report the polar cases $\beta_g = 0$ (fixed budgets: each additional student dilutes per-student spending one-for-one) and $\beta_g = 1$ (funding follows each child) alongside the estimated value.

Parameters measured outside the model. City productivities $\bar{A}_i$ are the 2005 wage per efficiency unit by tier (tier 2 normalized to one). The education budget share $s_E = 0.08$ is *public* education spending relative to labor income: national fiscal education expenditure reached approximately 4% of GDP by the early 2010s — the vintage of our household targets — which at a labor-income share of roughly one half amounts to $\approx 8\%$ of labor income. (The local

⁷Log monthly income on years of schooling, a quadratic in age, and gender, with census weights and prefecture-clustered standard errors; employed adults 16–60. Without the hukou restriction the all-worker ratio is 1.9, inflated by compositional selection into cities; among inter-county migrants only it is 1.1, attenuated by within-education selection and occupational downgrading. The composition-held 1.34 is the mapping the model object requires.

government-education line item alone, $\sim 3\%$ of GRP, understates the total because a large share of funding flows through central transfers.) Private household spending is deliberately excluded: it is the empirical counterpart of the model’s private input m and is disciplined separately by the education-expenditure-share moment. The birth-city distribution π_{birth} is the 2005 registered population by tier. The bilateral migration costs decompose as $d_{ii'} = \bar{d} \tilde{d}_{ii'}$: the relative structure $\tilde{d}_{ii'}$ is a 291×291 prefecture-pair *travel-time* matrix from [Ma and Tang \(2024\)](#), aggregated to the seven tiers by population-weighted averaging and normalized so that the average urban–urban cost equals one; diagonal entries are zero, and the rural row and column, for which prefecture pairs are not available, are set to uniform levels with the urban→rural cost 1.5 times the rural→urban cost. The scale $\bar{d}$ converts this structure into utility units and is calibrated jointly below. Table 3 reports the tier-level stage-one objects.

Table 3: City-tier objects set outside the model

	Tier						
	1 (rural)	2	3	4	5	6	7 (mega)
Urban population (million)	—	<0.5	0.5–1	1–3	3–5	5–10	>10
Productivity $\bar{A}_i$	0.33	1.00	1.04	1.22	1.50	1.90	2.01
Birth share π_{birth}	0.548	0.006	0.043	0.242	0.076	0.048	0.037
Enrollment prob. ρ_i	1.00 [†]	0.90*	0.88*	0.86*	0.75*	0.70*	0.57

Notes: Access schedule: the megacity value (tier 7) is a migrant-weighted average of published city-level enrollment shares; [†]the rural value is one by construction (schooling at the registration county faces no enrollment barrier); *assigned by interpolation between the observed city shares and the published national average, the schedule rescaled by 1.019 to match the aggregate (Appendix F).

4.2 Jointly calibrated parameters and targeted moments

Four parameters — the education scale z , the migration-cost scale $\bar{d}$, the settlement cost m^c , and the warm-glow weight η — are calibrated *jointly*: each candidate parameter vector requires solving the stationary equilibrium, and the vector is chosen so that the model matches four moments in levels. For the household-level targets we build a stacked person–wave panel from the CFPS adult questionnaires (2010, 2014, 2016, 2018), linking parents and children through the family-configuration files and applying CFPS sampling weights; the estimation sample is adults aged 16–60 with valid weights. The model has a single sector and no industry dimension, so all targets are household- or aggregate-level; we use external aggregates where CFPS coverage is weak. The four targeted moments are:

- **Education expenditure share of income.** Household education spending relative to labor income, from the CFPS family-economic module; disciplines the education-technology scale a jointly with the warm-glow weight η (under log utility the glow is additive in $\log h_c$, so the two act on spending through distinct margins — a through the return to inputs, η through the parent’s contemporaneous valuation of them).

- **Wage dispersion.** The raw standard deviation of log labor income among employed adults in the CFPS, 1.25 — education is deliberately *not* controlled for, since the model’s h embeds schooling, and the raw (rather than age-adjusted) dispersion is the appropriate reading since the enrollment lottery itself generates a structural component of income dispersion.
- **Rural-to-urban migration rate.** Among agricultural-hukou holders, the share currently residing in an urban area (census definition), measured in the CFPS 2014 wave: the survey begins only in 2010, and the hukou-based migrant stock is slow-moving, so this stock measure aligns with the post-liberalization vintage of the other household targets; disciplines the migration-cost scale $\bar{d}$.
- **Child co-residence rate.** Among children of rural–urban migrants, the share living with their parents at the destination rather than left behind: 0.37, the 2010-census rural-basis co-migration share (35.8 million migrant children against 61 million rural left-behind children); disciplines the settlement cost m^c . This is the complement of the widely cited “left-behind” share (the source of the 61-million count; [All-China Women’s Federation, 2013](#); [Zhang et al., 2014](#); [Meng and Yamauchi, 2017](#)), and its vintage matches the other household targets.

Table 4 collects all parameter values; Table 5 reports the fit.

4.3 Fit and identification checks

Two intergenerational moments are treated as checks rather than fitted targets. We match parent–child pairs through the CFPS father and mother identifiers and estimate an intergenerational elasticity by regressing a (log) child outcome on the (log) parent outcome, controlling for child age, age squared, and gender. Because income is observed for only a small number of matched pairs, we rely on the education-based elasticity (child years of schooling on parent years of schooling) for coverage, and compare the model’s implied elasticity of *human capital* to the external range of income-based estimates for China of 0.4–0.6 ([Fan et al., 2021](#)). Both checks pass: the model’s elasticity is 0.555, inside the external range, and the parent–child correlation is 0.593 against an empirical value near 0.5 (Table 5).

The level moments are matched essentially exactly (Table 5): expenditure share, migration, and co-residence land on target to three digits, and income dispersion overshoots the raw CFPS value by 6%. The overshoot is the visible cost of the σ choice of Section 4: a free σ is driven to zero, and a model in which child outcomes are deterministic in family inputs overstates persistence (the implied IGE rises to 0.69). At $\sigma = 0.30$ both intergenerational checks return to their empirical ranges, at the price of the 6% overshoot.

Table 4: Model parameters: values and identification

Par.	Description	Value	Identification / source
<i>Demographics and structure</i>			
J, J_c	Lifespan, childhood length (periods)	8, 3	one period = 7 years
I	Number of cities (rural + 6 urban classes)	7	2005 city-tier classification
<i>Predetermined — preferences and assets</i>			
β	Discount factor (per period)	0.70	annual ≈ 0.95
σ_u	Utility curvature ($u = \log c$)	1.0	scale-free migration choice
r	Risk-free return (per period)	0.15	annual $\approx 2\%$
<i>Predetermined — technology, mobility, policy</i>			
ϕ	Education money share	0.40	set a priori
ω	Education curvature	0.60	Ben–Porath curvature range (0.5–0.9)
δ	Human-capital spillover weight	0.30	HC-externality literature; set a priori
γ	Dynastic altruism weight	0.50	Barro–Becker range; collinear with z
κ	Gumbel scale, city choice	2.0	normalization
λ	Agglomeration elasticity (urban tiers)	0.06	Combes and Gobillon (2015) range
s_E	Education budget share (allocation rule)	0.08	public edu. spending / labor income
θ_u	Urban skill bias ($\theta_1 = 1$)	1.35	urban/rural returns-to-schooling gradient
β_g	Fiscal rule: budget elasticity to enrollment	0.51	own estimate (Fact 3, Table 2)
ψ	Conversion education-score slope	2.0	set (selection flat in ψ ; see text)
ρ_i	Status-quo access schedule (the policy instrument)	Table 3	enrollment shares, Chen et al. (2020); App. F
$\bar{p}_i$	Residence-conversion levels	data	hukou-conversion registration shares
$\bar{A}_i$	City productivities (tier-2 normalized)	data	2005 wage per efficiency unit
π_{birth}	Birth-city distribution	data	2005 registered population
$d_{ii'}$	Bilateral migration costs (scaled by $\bar{d}$)	data	travel times, Ma and Tang (2024)
κ_c	Bring-choice scale (settlement-cost dispersion)	0.50	benchmark value; robustness in text
σ	Std. dev. of the (log) education shock	0.30	disciplined by the income-IGE range (see text)
<i>Jointly calibrated (SMM)</i>			
z	Education scale	2.67	education expenditure share
$\bar{d}$	Migration-cost scale	8.34	rural→urban migration rate
m^c	Settlement cost of the child in the city	1.46	co-residence rate, census stocks
η	Warm-glow weight on child development	0.18	education expenditure share (with z)

Table 5: Model moments at the calibrated benchmark vs. data targets

Moment	Model	Target	Source
<i>Fitted (Stage-2 SMM)</i>			
Education expenditure share of income	0.120	0.120	CFPS family-economic module
SD(ln labor income)	1.33	1.25	CFPS, raw (σ set; see text)
Rural→urban migration rate	0.431	0.426	CFPS 2014 (agri-hukou urban)
Co-residence rate, rural migrants	0.375	0.370	2010 census stocks (rural basis)
<i>Over-identification checks</i>			
IGE (child–parent human capital)	0.555	0.4–0.6	income-IGE range (Fan et al., 2021)
Parent–child HC correlation	0.593	~ 0.5	same range
Urban share, top-skill-third rural-born	0.496	0.615	CFPS, top schooling third of rural-born
Excluded share among brought migrants	0.242	$1 - \bar{\rho} \approx 0.2\text{--}0.3$	implied by the benchmark ρ_i

Migration selection is an over-identification check earned by the anchored skill bias, and the model passes it where most of the population lives. Among rural-born adults the urban-residence share rises monotonically in skill, in the model (0.37, 0.43, 0.50 by thirds of the skill distribution) as in the CFPS, where it rises from 0.30 (primary or less) through 0.48 (junior high) to 0.65 (senior high) and 0.77 (college) — estimated on exactly the model’s population, rural-born adults whether or not they have since converted, and robust to age, gender, and childhood-province controls. The lower half of the gradient is matched closely. The one shortfall is at the top: the urban-residence share of the top third of rural-born adults is 0.615 in the data against 0.496 in the model (Table 5). This is the price of the anchoring: fitting θ_u to this moment would push it to 1.54 and stretch income dispersion to 1.68 against 1.25 in the data, so we accept the top-end shortfall rather than let one moment buy its own parameter. The excluded share among brought migrant children, 0.24, is consistent with the measured enrollment schedule by construction.

Cross-tier fit. Because every result in the paper is reported tier by tier, Table 6 compares the model’s benchmark spatial distribution — none of it targeted — with its census and city-panel counterparts. The model is closest where the results live, in the large cities: the three largest classes hold 52% of the model’s urban adults against 43% in the census, receive 67% of its rural-origin migrant parents against 71%, and carry a per-enrolled-child spending gradient of the same steep shape (tier 7 at 3.1 times tier 2 in the model, 3.9 in the data). The miss is in the small and middle classes: the model routes 20% of migrant parents to the two smallest classes against 7% in the census, at the expense of the largest middle class. The megacity incidence conclusions of Section 7 therefore rest on the best-matched part of the distribution, while the small-tier lineage’s gain from closure is computed where the model overstates the migrant inflow and should be read as illustrating the fiscal mechanism rather than as a measured stake.⁸

Mechanisms. Four further pieces of evidence, none targeted in the calibration, bear directly on the model’s working parts: two survey exhibits on the child states and the conversion margin, the model’s own account of the Fact-2 heterogeneity, and its counterpart of the 2014 event study. First, *the value gap behind the lottery*. The model’s child states have direct counterparts in the China Education Panel Survey, which administers a common cognitive test. Relative to rural children living with both parents, and conditional on gender, grade, and parents’ education (Table A5), migrant children *enrolled in public schools* score slightly above the reference group (+0.08, s.e. 0.04) and left-behind children at parity (+0.01, s.e. 0.02). Migrant children in non-public schools — the model’s excluded state — score 0.15 below it (s.e. 0.15). That level

⁸The rural–urban line itself is omitted from the comparison: the model’s tier 1 is the rural *registered* sector, while the census classifies by physical residence, and the two disagree precisely for rural-registered adults living in town areas of prefecture cities.

Table 6: Cross-tier fit: the model’s benchmark spatial distribution against the data

	Urban tier					
	2	3	4	5	6	7 (mega)
<i>Urban adults, share of urban total</i>						
model	0.055	0.095	0.327	0.172	0.178	0.174
census 2005	0.069	0.075	0.427	0.166	0.137	0.125
<i>Rural-origin migrant parents, destination shares</i>						
model	0.096	0.099	0.134	0.171	0.244	0.256
census 2005	0.037	0.032	0.225	0.155	0.263	0.289
<i>Per-enrolled-child spending, relative to tier 2</i>						
model (solved g_i^{nat})	1.00	0.65	0.70	1.24	2.39	3.13
city panel 2005	1.00	0.95	0.90	1.22	1.89	3.88

Notes: None of these distributions is targeted; the calibration’s only spatial moments are the tier-level productivities and the aggregate migration rate. Model: benchmark stationary equilibrium (adults of working age; migrant rows condition on rural-born child-rearing adults working in an urban tier). Census 2005: weighted; adults 21–56 by residence prefecture mapped to the six official size classes; migrant rows are agricultural-hukou inter-county migrants aged 21–42 residing in cities or towns. City panel 2005: education expenditure over primary-plus-secondary enrollment, aggregated by class. The rural tier is omitted because the model’s tier 1 is the rural *registered* sector while census residence classifies by location; the two cuts disagree for rural-registered residents of prefecture-city town areas.

estimate is too imprecise to stand alone because the excluded cell is thin: 228 students in seven non-public schools, reflecting the survey frame’s limited coverage of migrant schools. The object that maps to the model is the *within-migrant* enrolled-versus-excluded contrast, estimated directly: -0.20 (s.e. 0.07), about a quarter of a standard deviation. The enrolled–excluded gap that gives enrollment risk its stakes in the model is thus visible in measured skills, with the imprecision of the thin excluded cell as its caveat.⁹

Second, *the conversion gradient*. In the CFPS, the probability that a rural-born adult (agricultural hukou at age 3) holds non-agricultural hukou today rises from 5% among the illiterate to 16% with junior-high, 39% with senior-high, and 71–87% with college education ($n = 18,910$) — the steeply education-scored acquisition that the conversion technology $p(i, h)$ builds in, including its top end: for the educated, conversion approaches certainty, the behavior the specification’s product cap is designed to admit. The gradient survives birth-cohort, gender, and childhood-province controls (probit average marginal effects of +12, +31, and +73 percentage points for junior-high, senior-high, and college relative to primary or less). It also runs through the education channel itself: among adults still holding agricultural hukou at age 12, subsequent conversion rises from 6% to 75% across the same education groups — the college-admission

⁹CEPS baseline (2013–14), $n = 16,284$; standardized 3PL cognitive score. In value-added form (the wave-2 score conditional on the baseline score), the state gaps are statistically indistinguishable from zero over the survey’s one-year panel — a horizon too short to detect a gap accumulated over childhood, and one subject to differential attrition of migrant-school students — so we read the level gaps as cumulative exposure combined with selection, and use them as qualitative validation rather than a targeted moment.

route that makes residence acquisition education-scored.¹⁰

Third, *who responds to access*. The census interaction of Fact 2 (equation (1), Table 1) — family arrangements respond to enrollment restrictions three times more strongly for less-educated mothers — is reproduced by the model without being targeted. Removing enrollment risk ($\rho_i = 1$) raises co-residence by 5.5 percentage points for parents in the bottom third of the benchmark rural-born skill distribution, 3.6 points in the middle third, and 0.3 at the top (equal-population thirds; migrant parents' benchmark weights across them are 28, 33, and 38%). The aggregate long-run response of +5.1 points exceeds the weighted within-third average (+2.9) because the reform also operates on two margins the fixed-thirds comparison holds down: migration itself rises — the new migrant families co-reside at higher rates — and the rural-born skill distribution shifts up as better-schooled cohorts graduate. The self-insurance mechanism that caps the aggregate response also concentrates it on the least-skilled families, exactly as in the data.

Fourth, *the causal response to the 2014 episode*. The reform-based event study of Section 2 has a direct model counterpart: cut ρ in the two strictest tiers, hold prices frozen (the analogue of a short-horizon reduced-form design), and compute the same treated-minus-control change in destination-side separation. The model delivers +4.1 percentage points at a dose of $\Delta\rho = -0.15$ and +8.9 at -0.30 — roughly 28 points of separation per unit of enrollment probability — against the census estimate of +11.8. Read at the episode's face dose of -0.15 , the model delivers about a *third* of the estimated response; reproducing the estimate exactly requires an effective dose of $\Delta\rho \approx -0.4$ in the strict tiers, a severity consistent with what the points systems did to marginal migrant families' admission odds. Either reading makes the reform results conservative: if the reduced-form elasticity were taken at face value at the face dose, the equal-access reform's +5.1-point reunification response would scale to roughly +14 points — every reported result instead uses the model's own, smaller elasticity. The model's family-arrangement response to policy is thus the same order of magnitude as its quasi-experimental counterpart, without the response ever being targeted; the comparison is inexact in known directions (the model episode runs one seven-year period against the census's shorter exposure, and the census estimate includes migrant-composition shifts the frozen-price experiment only partly captures), so we read it as validation of scale, not a point match.

Computation. The equilibrium is solved as nested fixed points: the dynastic value function innermost, the within-lineage human-capital distribution next, and a tatonnement on the city-wage vector outermost; transitions are solved in the same way in time-indexed form, iterating a backward value sweep and a forward distribution sweep to convergence. Appendix I details the algorithm, the state space, and grid refinement. Two numerical points matter for the results.

¹⁰A lifetime stock, whereas the model's $p(i, h)$ is a per-period probability; the comparison disciplines the shape of the gradient, not its per-period level.

Self-productive human capital with a city-average spillover can in principle support multiple development equilibria (as in [Bénabou, 1996](#); [Durlauf, 1996](#)); at the calibrated parameters it does not — re-solving the benchmark and every reported counterfactual from initializations of $\bar{H}_i$ spanning the grid returns the same equilibrium in every case (Appendix I) — so no result depends on an equilibrium selection. And because the self-productive law yields a heavy right tail in human capital, the state grid is log-spaced: linear or too-narrow grids clamp the tail and artificially compress the dispersion and intergenerational moments.

5 Education access reform

The central policy experiment of the paper is the removal of enrollment barriers: an unanticipated, permanent reform at date $t = 0$ that sets $\rho_i = 1$ in every tier, starting from the size-declining benchmark schedule of Table 3. The reform thus removes a *gradient* of enrollment risk, not a uniform wedge: the enrollment probability of a brought migrant child rises by 0.10 in tier 2 and by 0.43 in the megacities, where under the status quo 43% of brought children were excluded from public school. Under the fiscal rule (7) the trade-off runs at the estimated pace: every migrant child now claims a full seat, and budgets absorb roughly half of the enrollment change ($\beta_g = 0.51$), so per-enrolled-child spending dilutes — partially — wherever the migrant-child load is large relative to natives, while the reform acts directly on the family-arrangement margin, since bringing a child no longer carries enrollment risk. Comparing the pre- and post-reform steady states would tell us how lineages fare in the distant long run, but nothing about the generations alive when the policy changes, whose lifetimes unfold along the adjustment path. This section therefore solves the perfect-foresight transition and traces the reform’s full incidence on the households alive at the reform — who gains, who loses, and through which channel — and closes by bracketing every result with the polar fiscal rules $\beta_g = 0$ (fixed local budgets) and $\beta_g = 1$ (funding follows the child), which separates what the reform does through access from what it does through financing. One model period is seven years; “on impact” therefore always means *within the first seven-year model period*, not on the reform date itself.

The pre-reform stationary distribution is inherited as a predetermined stock; from $t = 0$ on every household re-optimizes under the new policy with full foresight of the equilibrium paths of wages, per-enrolled-child spending, and the spillover. The tax-rate path of (8) is computed along the transition and charged against every household’s consumption, so all headline welfare changes are net of their experiment’s tax cost.

Figure 2 plots the paths. Panel (a) shows the dilution the fiscal rule permits — and reverses. Where the removed restriction was strict the dip is immediate but, at the estimated adjustment, shallow and temporary: per-enrolled-child spending in the megacities falls 8.7% on impact, reaches its trough of -12.5% one period later as the migration response builds, and

ends 9.2% *above* its pre-reform level in the long run, because the reform’s in-migration enlarges the megacity tax base (efficiency labor +27%) by more than its enrollment. Tiers 2–5, already near enrollment parity, barely dip (–0.2 to –3.4% at the trough) and end 10–28% higher; rural spending rises 28%. Dilution, under the estimated rule, is a *transition* phenomenon, but a long one: megacity spending per enrolled child stays below its pre-reform level for fourteen model periods — roughly a century — though the dip is deep only in the first two (below –10%) and shallower than –3% from the tenth on. The *welfare* cost is nonetheless concentrated on the cohorts alive at the reform (Table 9): later school cohorts meet a mild residual dip against steadily higher graduation human capital and city wages.

Panel (b) splits children by schooling location and shows the same compression of the reform’s characteristic dynamic: migrant children’s mean graduation human capital dips 1.6% on impact — previously-enrolled children gain nothing directly from the reform while the budget dip bites at once — and then climbs to +28% in the long run as previously-excluded children enter public school, reunified families accumulate, and the recovering budgets compound the gains. Native children gain in every tier in the long run — +3 to +10% in tiers 2–5 and, once budgets recover, +5.7% in tier 6 and +4.6% in the megacities, whose native children suffer only a transitory dip of about 2.4% — and rural children gain +17%.

Panel (c): the rural average human capital rises by roughly a third on a generational clock as better-educated rural-lineage graduates enter.

Panel (d) shows the reform acting on family arrangements and location choice at once: the co-residence rate of rural migrants jumps from 37.5% to 42.1% in the first period and settles at +5.1 percentage points, while the left-behind rate dips 0.9 points on impact and then returns essentially to its pre-reform level. The flat aggregate is the product of two margins the reform moves in opposite directions. The left-behind share of rural-born children is (share of parents who migrate) $\times$ (share of migrants who leave the child behind): before the reform, $0.431 \times 0.625 \approx 26.9\%$. After it, each migrant family is less likely to leave the child behind ($0.625 \rightarrow 0.574$), but migration itself rises ($0.431 \rightarrow 0.477$). Removing enrollment risk raises the value of migrating *with* children, so families who previously stayed rural now go, and some of the newcomers still leave the child behind (the settlement cost m^c is unchanged). The product is $0.477 \times 0.574 \approx 27.4\%$: the two effects offset almost exactly. The timing follows the same logic: on impact the stock of migrant parents is predetermined, so only the reunification margin operates and the left-behind rate falls; the migration response then builds over the following periods and refills the stock. Reunification thus takes the form of more migrant families co-residing rather than fewer migrants — the urban adult share rises by 2.6 percentage points — and the children who remain left behind attend rural schools whose spending is 28% higher. City wages still move by less than 2%: the reform’s action runs through what children learn and where families live, not through prices.

Table 7 collects the long-run destinations of these paths, tier by tier. The family-arrangement

and migrant-child columns are broadly monotone in the removed restriction $1 - \rho_i$; the spending column runs the *other* way — the long-run budget gains are largest where restriction was mildest, because the tax-base growth that finances them arrives everywhere while the enrollment expansion concentrates where exclusion was severe.

Table 7: Effects of removing enrollment barriers ($\rho_i \rightarrow 1$) by city-size tier, long run versus pre-reform steady state

Tier	ρ_i	Spending per enrolled child		Co-residence (destination)	Grad. HC, migrant children	Grad. HC, home-tier	Adult pop. share
		trough	long run				
1 (rural)	1.00	+0.1%	+27.9%	—	—	+16.9%	−2.6pp
2	0.90	−3.4%	+28.4%	+4.0pp	+21.6%	+9.3%	+0.1pp
3	0.88	−1.1%	+24.4%	+3.3pp	+20.5%	+10.1%	+0.1pp
4	0.86	−0.3%	+10.3%	+1.8pp	+17.0%	+3.5%	−0.0pp
5	0.75	−2.8%	+17.1%	+4.5pp	+22.0%	+6.2%	+0.3pp
6	0.70	−6.9%	+14.6%	+5.1pp	+25.7%	+5.7%	+0.8pp
7 (mega)	0.57	−12.5%	+9.2%	+7.7pp	+33.7%	+4.6%	+1.3pp

Notes: Spending per enrolled child: the transition trough and the new steady state, relative to the pre-reform steady state. Co-residence: among migrant parents working in the tier. Grad. HC, migrant children: by schooling destination; home-tier: children schooled in their home tier (for tier 1, all rural children including the left-behind). On the origin side, the left-behind rate of rural-born children is essentially unchanged (26.9% to 27.4%); co-residence rises among a migrant population that itself grows. Urban-origin left-behind rates are small ($\leq 13.5\%$ and declining in tier) and essentially unchanged. City wages move by less than 2% throughout.

Decomposition: where the effects come from. Table 8 splits the impact effects into three layers: the *mechanical* effect of the new enrollment odds (families keep their pre-reform decision rules and values, prices frozen), the *behavioral* response (values and policies re-solved at frozen prices — a partial-equilibrium counterfactual matching the short-run design of the reduced-form evidence), and the *full equilibrium*. The family-arrangement response operates already in the mechanical layer: better odds raise the value of bringing a child at pre-reform decision rules, lifting rural-migrant co-residence by +5.8 percentage points in the first period; re-optimization trims this to +5.2, and general-equilibrium feedback — the transitory dilution of the very budgets access was granted to — to the observed +4.6. The *impact dip* in migrant children’s human capital is, by contrast, entirely general-equilibrium: the mechanical and frozen-price effects are −0.0% and −0.1%; the budget dip makes it −1.6%. Parental input crowd-out, a large force under a spending-discount formulation, is absent outright — in full equilibrium parental money and time inputs to migrant children *rise* on impact: previously-enrolled children never gain spending they could offset, and the warm glow keeps parents investing in the newly enrolled. On the welfare side the split is sharp. Migrant families’ gains are partial-equilibrium at their core, with general equilibrium clawing back about a fifth: the megacity migrant gain is $+174 \times 10^{-3}$ at frozen prices and +136 in full equilibrium. Urban natives’ losses are *entirely* general-equilibrium — their frozen-price change is ≈ 0 . What a

native family loses from the reform is not access, which it always had, but the transition-period budget dip; that is why the estimated fiscal rule matters so much for their side of the ledger. The rural children’s long-run gain is mostly compositional (better-educated lineages), with the equilibrium rise in rural spending and $\bar{H}$ adding the remainder.

Table 8: Decomposition of the reform’s impact effects: mechanical odds, behavioral response, and general equilibrium

	mechanical (odds only)	+ behavior (frozen prices)	+ prices (full GE)
Co-residence, rural migrants ($t=1$, pp)	+5.8	+5.2	+4.6
Left-behind rate ($t=1$, pp)	−1.5	−1.0	−0.9
Grad. HC, migrant children ($t=0$, %)	−0.0	−0.1	−1.6
ΔV , migrant families, megacity (10^{-3})	—	+174	+136
ΔV , native families, megacity (10^{-3})	—	≈ 0	−64

Notes: Mechanical: the new ρ_i applied to pre-reform values, decision rules, and prices — the direct effect of better odds on the enrollment draws and the odds-weighted bring choice. Behavioral: values and policies re-solved at prices frozen at the pre-reform steady state. Full GE: the transition of Figure 2, under the estimated fiscal rule $\beta_g = 0.51$. The mechanical layer implies no value changes, so ΔV is defined only for the last two columns.

In the aggregate, the reform is a human-capital event that prices barely register: the cohort-average graduation stock rises 17% in the long run while wages move by less than two percent, because the labor force absorbs improved cohorts one period at a time and $\lambda = 0.06$ prices what remains. Its first-order consequences are therefore distributional — across groups, across tiers, and across family arrangements — and the sections that follow trace them.

We score each household alive at the reform by the change in its value relative to counterfactually remaining in the old steady state, $\Delta V = V_{t=0}^{\text{trans}} - V^{\text{old SS}}$, and aggregate the mean change and the *winner share* (mass with $\Delta V > 0$) within cells defined by tier, native-versus-migrant status, and life stage. Alongside utility units we report each change as a *consumption-equivalent variation* (CEV): the permanent proportional change in the dynasty’s consumption path that would produce the same value change. Under log utility the conversion is exact and common to all states. A permanent proportional increase x shifts every value by $\log(1+x) D$, where $D = (\sum_{a=0}^{J-J_c-1} \beta^a) / (1 - \gamma\beta^{J_c}) = 3.35$ is the discounted dynastic count of consumption periods; additive location shocks, migration costs, and the warm glow (which does not scale with consumption) all pass through unchanged. Hence $\text{CEV} = \exp(\Delta V/D) - 1$.¹¹ One accounting choice deserves emphasis. Measured household welfare includes the warm-glow flow — the same object that drives the family-arrangement behavior — and counting it in social welfare could be seen as letting the mechanism grade itself. We therefore evaluate every steady-state comparison under two criteria: household preference welfare W , and a *social criterion* that

¹¹For adults already partway through life the exact constant differs by less than 1% at every age with a dependent child (the dynastic continuation dominates D); we use the newborn-dynasty constant throughout.

counts consumption and the child’s future utility but strips the paternalistic flow,

$$W^S(i, h) = W(i, h) - \mathbb{E} \left[\sum_{g=0}^{\infty} (\gamma \beta^{J_c})^g \sum_{j=0}^{J_c-1} \beta^j \eta \log \bar{h}'_c(g, j) \right], \quad (9)$$

where $\bar{h}'_c(g, j)$ is the child human-capital accrual of the dynasty’s generation- g parent at parenting age j , and the expectation runs over the dynasty’s equilibrium path: W^S subtracts the discounted dynastic present value of the warm-glow flows at *unchanged* behavior, so the two criteria differ in accounting only. The two never disagree on a ranking, and roughly 90% of every reform’s gain survives the stripping (Table 12): the gains are consumption and dynastic continuation, not the glow term valuing itself. Readers who prefer the non-paternalistic criterion as the baseline can read every conclusion from the glow-excluded rows. Table 9 reports the result.

Table 9: Winners and losers at the reform date: value changes, consumption equivalents, and winner shares by tier and group

Tier	Native families			Migrant families			Entrants (by lineage)		
	ΔV	CEV	win%	ΔV	CEV	win%	ΔV	CEV	win%
1 (rural)	+53	+1.6%	100	−26	−0.8%	64	+119	+3.6%	100
2	−24	−0.7%	0	+31	+0.9%	93	−15	−0.4%	100
3	−13	−0.4%	100	+35	+1.1%	93	+0	+0.0%	100
4	−14	−0.4%	100	+39	+1.2%	93	−3	−0.1%	100
5	−25	−0.7%	0	+57	+1.7%	95	−24	−0.7%	0
6	−48	−1.4%	0	+78	+2.4%	80	−52	−1.5%	0
7	−81	−2.4%	0	+119	+3.6%	79	−89	−2.6%	0

Entrants averaged over the stationary entrant skill distribution (same object, same net-of-charge basis, tiers 1–7): +85, −19, −9, −11, −23, −45, −77 $\times 10^{-3}$; allocation-basis winner shares 100/0/100/100/0/0/0.

Notes: “Families” are with-child adults (ages J_c+1 to $2J_c$), native if resident in their birth tier and migrant otherwise; cell means weight every state in the cell (age, skill, wealth, child status) by the pre-reform stationary mass. Entrants are the model’s newborn adults — agents beginning working life exactly at the reform date, the same object whose steady-state lineage values appear in Table 12 — weighted by π_{birth} , grouped by birth lineage, and evaluated at the lowest entry-skill node (a reference-newborn convention). ΔV is the mean value change in 10^{-3} utils; CEV is its dynastic consumption-equivalent, $\exp(\Delta V/D) - 1$ with $D = 3.35$. win% is the mass share of the cell with $\Delta V > 0$: 0 (100) means the cell loses (gains) at *every* state — native families’ exposure runs through a single tier-wide channel, the transition budget dip, so their sign is uniform within a tier — whereas migrant families are heterogeneous in exposure. No-child adults (not shown) have $\Delta V \approx 0$ everywhere. Cell values are *net* of the reform’s common tax charge: the dynastic cost of its tax path, -17×10^{-3} , is applied additively to every cell, so the body follows the paper’s net-of-financing convention (allocation-basis values are recovered by adding 17). win% is computed at the allocation, before the common charge, so it isolates each cell’s own exposure channel: the tier-3/4 native and tier-2–4 entrant cells gain at the allocation (win% = 100) but show small net losses once the financing charge is applied. The last line reports the entrant object averaged over the stationary entrant skill distribution: the reform’s value falls in skill, so the reference evaluation is the upper end for gaining lineages and the lower end for losing ones; the steady-state tables use the distribution average throughout. Computational details are in Appendix I. Values include the warm-glow flow terms; differences are comparable within the model.

The reform’s incidence is sharply *heterogeneous* — and larger than a spending-discount formulation would suggest: a quarter more aggregate welfare and half again the family-arrangement

response (Table 11) — and it tracks the severity of the removed restriction. Two kinds of objects appear throughout. *Native* and *migrant* families are living adults at the reform date, classified by whether they reside in their birth tier; a *lineage* is the dynasty itself, indexed by birth city and evaluated at labor-market entry, the object of the steady-state comparisons. The same family appears in both ledgers; the ledgers answer different questions — how the reform treats the people alive when it arrives, and how it treats the bloodline viewed from birth. All values here are net of the reform’s common tax charge (the dynastic cost of its tax path, -17×10^{-3} , borne by every cell). *Migrant families gain on average in every urban tier*, from $+31 \times 10^{-3}$ (a 0.9% consumption equivalent) in tier 2, where enrollment was already near-certain, to $+119 \times 10^{-3}$ (+3.6%) in the megacities, where the reform removes a 43% exclusion risk; within the group 5–21% still lose, and the small group of migrants working in the rural tier loses (-26×10^{-3}), as rural spending is shared with in-migrating children and the financing charge lands on top. *Native families’ losses line up with the transition dip in the budgets that shield them* — and, under the estimated fiscal rule, they are losses of the transition, not of the new world. The megacity loss is -81×10^{-3} (–2.4% of consumption): megacity school budgets dip 12.5% at the trough while every migrant child claims a full seat. Tier 6 loses –48 and tier 5 –25, against –13 to –24 in tiers 2–4, where the restriction barely bound and the loss is essentially the common financing charge itself (at the allocation, before the charge, tier-3/4 natives gain at every state). The gradient sits where the *policy* was strict, tiers 5–7. Rural (tier-1) natives gain ($+53 \times 10^{-3}$) — their schools improve as spending drifts up and educated lineages accumulate — and *rural-born entrants gain most of all* ($+119 \times 10^{-3}$, a 3.6% consumption equivalent, and they are 55% of all births): their children, the migrants of the next generation, now face no enrollment risk wherever they go. Among urban entrants the reform-date pattern runs down the hierarchy: lineages born in tiers 2–4 start approximately whole (0 to –15), while entrants born into tiers 5–7 start behind (–24, –52, –89) — they eat the budget dip through their own children’s schooling years, while the recovered budgets and the compounding human-capital gains accrue to cohorts born later. That wages barely move while family arrangements move substantially confirms the operative channels: schooling access and family reunification, not factor prices.

The long-run comparison inverts the top of that ledger, and the estimated fiscal rule is what does the inverting. In the steady states (Table 12), equal access raises entrant welfare by 28% of permanent consumption for the average dynasty, net of its tax cost ($\Delta W = +0.84$), and *every* lineage gains: the rural-born most (+47%; $\Delta W = +1.29$), and the megacity-born too (+8.2%; +0.26). The reason is that by the long run budgets have absorbed the enrollment change, in-migration has enlarged the urban tax bases, and the financing tax (8.6% against the 8% anchor) is paid by everyone — so even the most protected lineage prefers the open regime. The intergenerational elasticity falls from 0.555 to 0.433. The reform’s redistribution is therefore *intertemporal* rather than dynastic: the megacity families alive at the reform pay (–2.4%

of consumption), and every lineage collects in the long run. Read in reverse, the status quo is *regressive* — because most children are born rural, the restriction transfers welfare from the rural majority, whose children are the migrants, to the urban natives it shields from the transition — and, under population weights, *inefficient*: the migrant human-capital loss, amplified by agglomeration and the spillover, far outweighs the transition dip it spares the incumbents. This is the human-capital counterpart of the regressivity that [Imbert et al. \(2023\)](#) document for Chinese migration policy.

Access versus financing: the fiscal bracket. The equal-access experiment bundles two policies — who may enroll, and who pays for it — and the fiscal rule is where the second is set. Table 10 therefore reruns the entire transition under the two polar rules, holding the access reform fixed: $\beta_g = 0$, fixed local budgets, in which incumbents bear the full dilution; and $\beta_g = 1$, funding that follows each newly enrolled child — the reform paired with full per-pupil financing, as if a central government topped up every seat. Because the benchmark is invariant to β_g by construction, the three columns share identical pre-reform economies and isolate the financing assumption exactly. The bracket makes three points. First, *the access reform is valuable under any financing*: even with rigid budgets the reform delivers +15% of consumption ($\Delta W = +0.46$). Second, *financing determines who pays*. Under fixed budgets, megacity spending never recovers (−21% in the long run) and the reform permanently taxes the top lineages (megacity-born −10% of consumption at entry; $\Delta W = -0.36$). Under the estimated rule, the dip is transitory, the tax rate settles 0.6 percentage points above the anchor, and every lineage gains. Under full per-pupil funding there is no dilution at all, but its financing — a tax rate of 9.8% against the 8% anchor — is charged to everyone. So the reform is not a free lunch even there: megacity natives at the reform date lose 0.4% of consumption ($\Delta V = -14 \times 10^{-3}$), a tenth of their fixed-budget loss but not a gain. Third, *the financing margin is itself a large policy*: moving from the estimated rule to funding-follows-the-child roughly doubles the welfare gain (from a 28% to a 68% consumption equivalent net of the tax cost; $\Delta W +0.84$ to $+1.74$) and raises migrant children’s long-run human-capital gain from +28% to +68%. The experiments of the following sections use the estimated rule; where the polar rules change a conclusion, we say so.

The estimation range. The polar cases bound the mechanism; the instrumented estimates of Appendix E.8 bound the *parameter*, and the reform’s conclusions hold across that range. Resolving the reform at the demographic-IV point estimate, $\beta_g = 0.30$, and at the upper confidence bound of the panel estimate, $\beta_g = 0.59$, brackets the headline gain between +22% and +32% of consumption net of the tax cost ($\Delta W +0.66$ and $+0.93$, against $+0.84$ at the estimated value); the rural-born gain runs from +37% to +52% (+1.06 to +1.41), reunification from +4.3 to +5.5 percentage points, and the fall in the IGE from 0.11 to 0.13. The one property

with a threshold inside the range is lineage-Pareto dominance: the megacity-born lineage's long-run gain shrinks from +13% of consumption ($\Delta W = +0.40$) at $\beta_g = 0.59$ to +8.2% (+0.26) at the estimated value and crosses zero near $\beta_g \approx \frac{1}{3}$ (−1.4%, $\Delta W = -0.05$, at the IV point) — continuously approaching the fixed-budget case, in which the top lineages pay for the reform (−10%; −0.36). The paper's aggregate, distributional, and mobility conclusions are insensitive to where in the credible range β_g lies; whether the very top lineage is a net gainer or approximately unharmed is the one margin the elasticity decides. The same is true of the rule's wage-bill exponent (Table A17):¹² at the estimated fiscal-expenditure elasticity in place of the imposed unit value, megacity budgets settle below rather than above their pre-reform level and the megacity-born lineage's small gain becomes a small loss, with every other conclusion intact. The policy instrument's near-zero point estimate — an impact response to a single shock, with a wide confidence band (Appendix E.8) — selects, at its most literal, the fixed-budget column of Table 10, already reported: a world of *stronger* dilution than the baseline, in which the reform keeps most of its value and incumbents bear its cost.

Table 10: The access reform ($\rho_i \rightarrow 1$) under three fiscal rules: fixed budgets, the estimated rule, and funding-follows-the-child

	$\beta_g = 0$ fixed budgets	$\beta_g = 0.51$ estimated	$\beta_g = 1$ follows child
Megacity g (impact/trough/long, %)	−17/ − 24/ − 21	−8.7/ − 12.5/ + 9.2	0/0/ + 60
Grad. HC, migrant children (impact/long, %)	−3.2/ + 13	−1.6/ + 28	−0.2/ + 68
Co-residence, rural migrants (long, pp)	+3.4	+5.1	+9.1
Tax rate τ , long run (%; anchor 8.0)	8.0	8.6	9.8
Δ entrant welfare (CEV)	+0.46 (+15%)	+0.84 (+28%)	+1.74 (+68%)
ΔW_7 , megacity-born (CEV)	−0.36 (−10%)	+0.26 (+8.2%)	+1.45 (+54%)
ΔV , megacity natives at reform (CEV)	−140 (−4.1%)	−81 (−2.4%)	−14 (−0.4%)
IGE under reform (benchmark 0.555)	0.464	0.433	0.386

Notes: Each column solves the full perfect-foresight transition and the terminal steady state under the indicated fiscal rule, with the access reform identical across columns. All welfare rows are net of the tax charge: the dynastic cost of the column's tax path (8), common to every household (−23 $\times 10^{-3}$ long run and −17 at the reform date under the estimated rule; −66 and −39 under funding-follows-the-child; zero under fixed budgets, where the rule spends exactly the anchor revenue). The ΔV row is in 10^{-3} utils. The benchmark equilibrium is invariant to β_g (at the benchmark $K_i = \bar{K}_i$), so all columns start from the same economy. This table is the separation the reform bundle implies: column 1 is equal legal access with no additional funding; column 3 is equal access with full central per-pupil funding.

The lottery against its certainty-equivalent. The claim that the lottery formulation matters is itself testable in the model, and Table 11 tests it: we re-solve the entire economy under the certainty-equivalent shortcut — a brought migrant child is always schooled and simply receives

¹²The fiscal rule scales each city's education budget one-for-one with its local wage bill $w_i N_i$ — the local tax base — an exponent of one that is imposed, not estimated. The alternative sets this exponent to 0.55, the elasticity of city education spending with respect to total fiscal expenditure in the Fact-3 specification (Table 2), so only about half of the tax-base growth that reform-induced in-migration delivers passes through to school budgets.

the discounted public input $\rho_i g_i$ — holding every parameter at its benchmark value. Three differences carry the comparison. First, the discount economy has *no excluded state*: every brought child is enrolled on partial funding, against the 24% of brought migrant children outside public school in the benchmark and the data — the institution’s observable footprint, with its quarter-standard-deviation skill gap (Table A5), cannot exist under the discount. Second, the lottery amplifies the family-arrangement response by half. The margin responds under both formulations, because the discount still prices access into the bring decision through the spending a brought child receives. But in the model counterpart of the 2014 episode, destination-side separation moves +4.1 percentage points under the lottery against +2.9 under the discount — a ratio of 1.4. The ratio rises to 1.7 when the bring choice sharpens ($\kappa_c = 0.25$), and it is unchanged on a doubled wealth grid. The amplification therefore comes from the uninsurable discrete risk itself — exclusion hits the family budget in the period it occurs — not from a grid artifact. Both formulations are compared under the model’s no-borrowing asset structure, which is what makes the within-period budget shock bind. Third, the stakes of reform are larger under the lottery: equal access is worth a 29% consumption equivalent in entrant welfare from the lottery benchmark against 23% from the discount benchmark (both before financing; $\Delta W = +0.86$ against +0.69), and the reform’s reunification response is +5.1 points against +3.4. The two formulations coincide at $\rho_i = 1$, so the reform steady state is common; what the discount misses is how costly the status quo is, and for whom.

Table 11: The enrollment lottery against the certainty-equivalent spending discount

	lottery (benchmark)	discount ($\rho_i g_i$ certain)
<i>Benchmark economy</i>		
excluded share, brought migrant children	0.24	0 (by construction)
co-residence rate, rural migrants	0.375	0.392
intergenerational elasticity	0.555	0.526
<i>2014-style episode: separation DiD (pp)</i>		
benchmark ($\kappa_c = 0.5$)	+4.1	+2.9
$\kappa_c = 0.25$	+4.6	+2.7
$\kappa_c = 1.0$	+2.5	+1.9
doubled wealth grid ($n_y = 8$)	+4.1	+2.8
<i>Equal-access reform ($\rho_i \rightarrow 1$), steady state</i>		
co-residence response (pp)	+5.1	+3.4
entrant welfare gain ΔW (before financing)	+0.86	+0.69

Notes: The discount column re-solves the full model with the enrollment draw replaced by certain schooling at spending $\rho_i g_i$ for brought unconverted migrant children, all parameters at benchmark values; the education budget is unchanged (expected outlay is identical). Episode: ρ cut by 0.15 in tiers 6–7 at frozen prices, destination-side separation change in the treated tiers net of tiers 2–4, as in Section 4.3; the κ_c rows vary the bring-choice scale holding all other parameters fixed in both columns. Reform rows compare each formulation’s own benchmark with the common $\rho_i = 1$ steady state (the formulations coincide at $\rho_i = 1$); ΔW is the population-weighted entrant welfare change before the financing charge.

Robustness. Appendix [H](#) collects four batteries of robustness checks; none changes a conclusion. First, the enrollment process: re-solved with admission decided once per school career instead of drawn i.i.d., the policy elasticity is essentially unchanged (+4.2 against +4.1 points) and the reform’s gains are slightly smaller. Second, the warm glow: with $\eta = 0$ the episode response is still +3.6 points — the family-arrangement margin needs the enrollment risk, not the paternalistic term. Third, the amplifiers: shutting down agglomeration, the human-capital spillover, or half the altruism one at a time leaves the reform large, lineage-Pareto, and mobility-improving (Table [A16](#)); removing the spillover actually *raises* the gain, so the result rests on the schooling channel. Fourth, the scale-setting assumptions — the excluded-state input, the conversion wage, congestion, the budget base, and the wage-bill elasticity — move the gain’s size but not its sign (Table [A17](#)); as with the enrollment elasticity’s lower range, the wage-bill elasticity decides only the sign of the megacity lineage’s small stake.

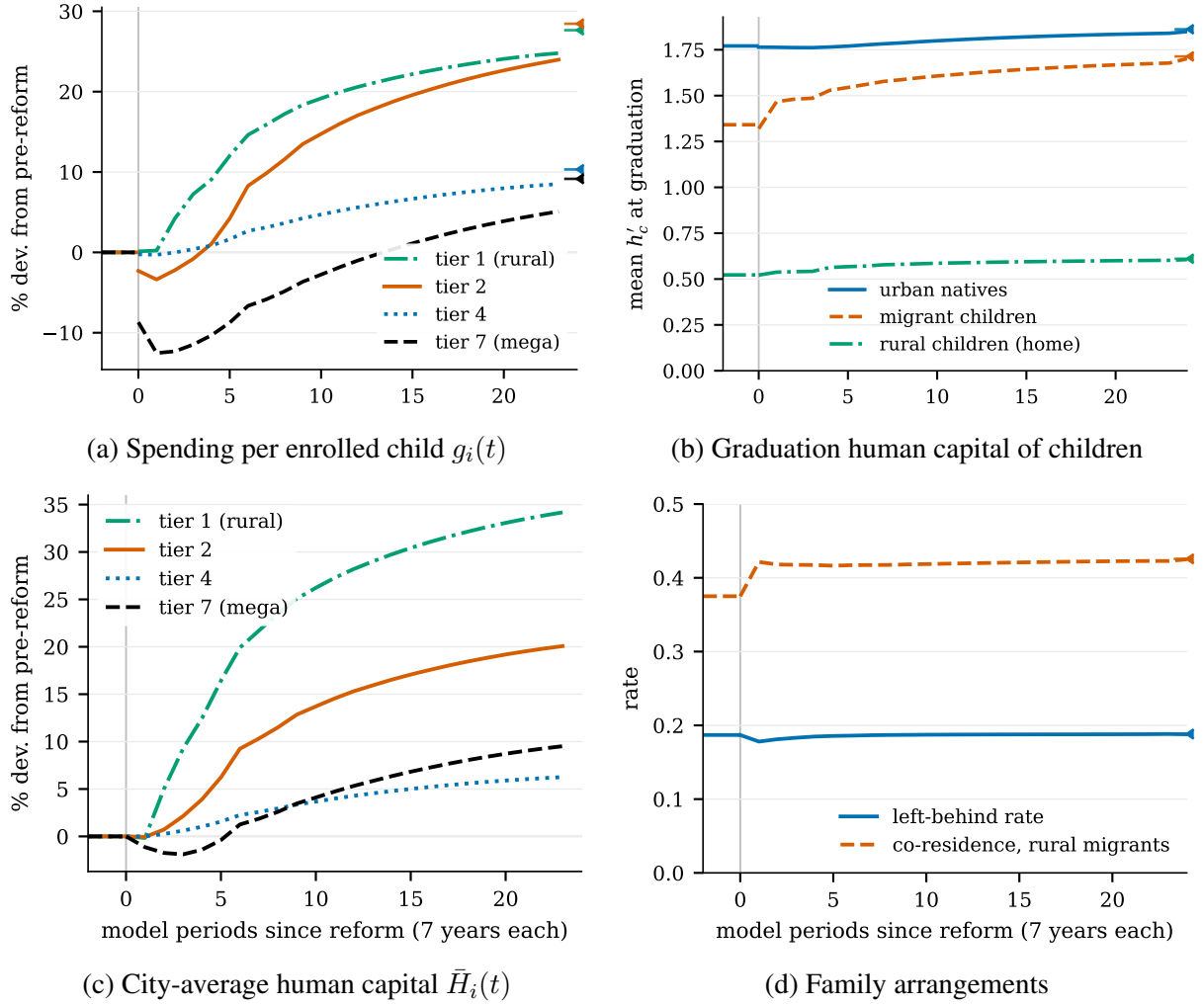


Figure 2: Transition paths after the equal-access reform

Notes: The benchmark schedule of Table 3 moves to $\rho_i = 1$ at $t = 0$; one period is seven years. Panels (a) and (c) are percent deviations from the pre-reform steady state; tiers 3, 5, and 6, omitted for readability, lie between the plotted tier-4 and tier-7 paths. Panel (b) shows the mean graduation human capital of children by schooling location: urban natives and rural children schooled in their home city (the latter including left-behind children), and migrant children schooled away from home; panel (d) the left-behind rate (share of dependent children living apart from their parent) and the co-residence rate of rural migrants. The vertical line marks the reform; the flat segments to its left are pre-reform values. Right-edge ticks mark the *separately solved* post-reform steady state — the long-run values the tables report; at the plotted horizon the slow-moving states (the distribution and city human capital) are still converging toward it (Appendix I reports the terminal residuals).

6 Alternative policy experiments

The transition shows access reform acting on both schooling and family arrangements; this section compares it with the other two instruments of the hukou menu introduced in Section 3: universal residence conversion ($p(i, h) = 1$ — every migrant family acquires local status on arrival) and a cut in the settlement cost of keeping a child in the city (m^c reduced by 20%). Each instrument is evaluated *alone*, holding the other two at their benchmark values; we defer combinations of instruments to future work. One caveat governs the comparison: the instruments are not cost-comparable. Access and conversion are administrative entitlements whose fiscal consequences the model prices through the budget rule, but the settlement cut reduces a utility friction whose real-world counterpart — family housing, childcare, documentation simplification — would carry resource costs outside the model. The comparison therefore ranks *what each margin buys*, not policies per unit of expenditure. Table 12 reports the new steady states, and Figure 3 the transition dynamics of each reform.

Table 12: The three hukou instruments, one at a time: steady states under access reform ($\rho_i = 1$), universal residence conversion ($p = 1$), and a settlement-cost cut ($m^c - 20\%$)

	benchmark	access $\rho_i = 1$	conversion $p = 1$	settlement $m^c - 20\%$
Co-residence rate, rural migrants	0.375	0.426	0.806	0.457
Left-behind share (all children)	0.187	0.188	0.059	0.177
Rural→urban migration rate	0.431	0.477	0.494	0.467
Excluded share, brought migrants	0.242	0.000	0.005	0.241
Mean log child graduation HC	0.010	0.209	0.511	0.125
IGE (child–parent human capital)	0.555	0.433	0.381	0.513
Tax rate τ , long run (%; anchor 8.0)	8.0	8.6	9.7	8.2
Entrant welfare, π_{birth} -weighted	−0.138	0.724	1.871	0.386
Δ vs. benchmark, net of the tax cost (CEV)	—	+0.84 (+28%)	+1.95 (+79%)	+0.52 (+17%)
Δ , warm glow excluded	—	+0.77	+1.72	+0.47
rural-born (W_1)	−3.549	−2.235	−0.660	−2.750
tier-2-born (W_2)	2.625	3.180	3.747	2.909
megacity-born (W_7)	7.287	7.573	7.919	7.537

Notes: Steady states under the estimated fiscal rule ($\beta_g = 0.51$). Each column changes one instrument from the benchmark, holding the other two fixed. Conversion at $p = 1$: every migrant family with a child converts on arrival in the work city (the $\bar{p}_i$ levels set to one, $\psi = 0$); the small residual excluded share belongs to families that have not moved since labor-market entry, as conversion is drawn on arrival. W_i is the value of a labor-market entrant of lineage i in the stationary equilibrium, averaged over the stationary entrant skill distribution. CEV converts the welfare difference to a permanent dynastic consumption equivalent, $\exp(\Delta W/D) - 1$, $D = 3.35$. Net of the tax cost: the Δ rows subtract the dynastic cost of each experiment’s tax path (8) (−0.023, −0.061, −0.008); lineage levels W_i are shown at the allocation, and the common charge does not affect comparisons across lineages within a column. The warm-glow-excluded row evaluates the same equilibria under the social criterion that counts consumption and the child’s future utility but strips the paternalistic flow: no ranking changes. Welfare levels include the warm-glow flow terms; differences are the meaningful objects.

Three conclusions. First, *access reform and a settlement cut work through different margins, and the fiscal rule prices the difference*. Access operates on schooling risk — the excluded state

disappears, mean log graduation human capital rises by 0.20,¹³ and the IGE falls from 0.555 to 0.433 — while the settlement cut operates on family arrangements: co-residence rises +8.2 percentage points (against access's +5.1) and the left-behind share falls by a point, where access leaves it unchanged. In welfare terms access now clearly dominates (+28% against +17% of consumption, both net of the tax cost; ΔW +0.84 against +0.52): under the estimated fiscal rule the enrollment expansion partly finances itself through the tax base its migration response builds, which is exactly the margin a settlement cut does not touch.

Second, *universal residence conversion is the deep reform*. It bundles certain enrollment with the end of the ongoing settlement cost, and delivers more than twice the welfare gain of the access reform: a 79% permanent-consumption equivalent net of its 9.7% tax rate (ΔW = +1.95). Co-residence reaches 0.81 and the left-behind share falls to 5.9% — reunification is built into converted status, since the child follows the parent. The schooling and mobility gains are the largest of the menu (mean log graduation human capital +0.50, IGE 0.381). And under the estimated rule the migration response is the largest as well (0.494): certain status makes the ascent worth more, and the recovering budgets no longer punish the families it pulls in.

Third, *in the long run every instrument is lineage-Pareto — the instruments differ in who pays along the way*. Under the estimated fiscal rule, all three reforms raise the entry welfare of *every* lineage net of the tax cost: the rural-born always gain most (+47% of consumption under access, +133% under conversion, +27% under settlement; ΔW +1.29, +2.83, +0.79), and even the megacity-born — the lineage the restrictions protect — gains under each (+8.2%, +19%, +7.4%; +0.26, +0.57, +0.24). As with access itself, the top lineage is the margin the fiscal elasticity decides: below $\beta_g \approx \frac{1}{3}$ the megacity-born gain from access crosses zero (the estimation-range paragraph of Section 5 reports the bracket). The steady-state comparison therefore contains no dynastic constituency for restriction; what separates the instruments is the *transition incidence* documented below and the mobility ranking: conversion and access, in that order, are the instruments that raise intergenerational mobility the most, while the settlement cut, which moves families without touching enrollment risk, barely moves it (0.513).

Figure 3 traces the three reforms along their perfect-foresight transitions, and the paths sharpen the steady-state ranking in two ways. First, *the family margin moves at reform speed while the schooling margin moves at generational speed*. Conversion reunifies at once — co-residence jumps from 37.5% to 82% and the left-behind rate from 18.7% to 5% within a single period — but migrant children's graduation human capital reaches its +60% only over several generations, and megacity per-child spending falls 27% at the trough (before recovering to +9%) as nearly every migrant child claims a native seat at once. The access path is the same

¹³The log-mean gain exceeds what the 17% rise in the arithmetic mean (Section 5) alone would imply because access reform also *compresses the left tail*: the excluded state, which holds 24% of brought migrant children at the benchmark, disappears at $\rho = 1$, and the SD of log graduation human capital falls from 0.70 to 0.64. By the lognormal identity $E[\log h] = \log E[h] - \frac{1}{2}\text{Var}(\log h)$, the mean contributes 0.158 and the variance reduction 0.043.

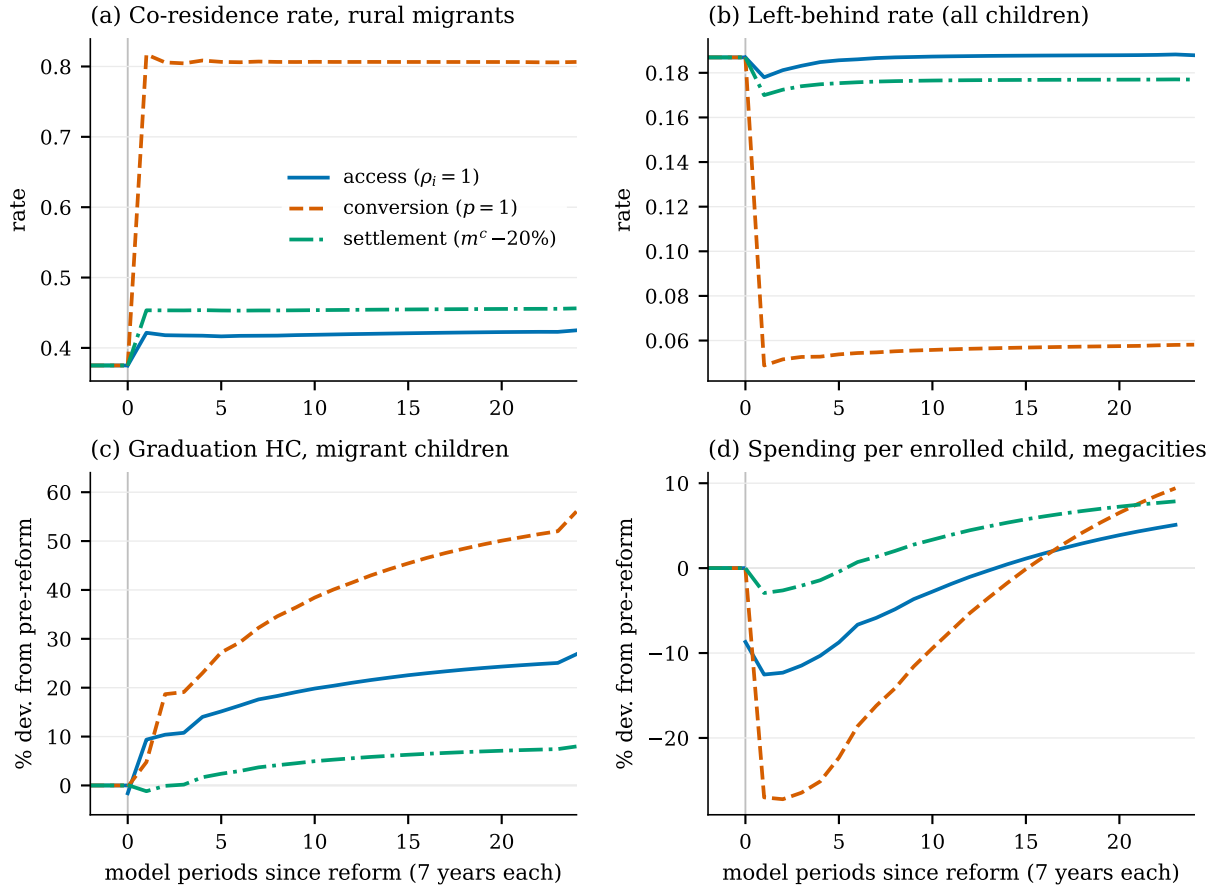


Figure 3: Transition dynamics of the three hukou instruments, each reformed alone at $t = 0$: access ($\rho_i = 1$), universal conversion ($p = 1$), and a settlement-cost cut ($m^c - 20\%$)

Notes: Panels (a)–(b) are levels; panels (c)–(d) percent deviations from the pre-reform steady state. The flat segments left of the vertical line are pre-reform values. One period is seven years.

shape at a smaller dose — megacity spending dips 8.7–12.5% and carries the impact dip of Section 5 — and the settlement cut is nearly transition-free: reunification is immediate and permanent (+8 points), spending barely moves (−3% at the trough), and there is no dip, because the children it brings enroll under the unchanged lottery rather than claiming full seats.

Second, the timing gap reverses conversion’s welfare verdict for the generations alive at the reform. Scoring the living as in Table 9, universal conversion makes losers of most of urban China alive at $t = 0$. Native families lose in every tier (−37 to -181×10^{-3}), urban-born entrants lose everywhere (−72 to −291), and migrant families themselves lose in tiers 2–6 (−33 to −133): the deep but transitory dilution of the schools their children just entered arrives within their lifetimes, while the human-capital gains accrue to cohorts not yet born. Only rural natives, rural-born entrants, and (barely) megacity migrant families (+8) gain on impact. The deepest steady-state reform thus has the harshest transition, and the mild settlement cut the gentlest — under the estimated rule the settlement cut is nearly painless for the living everywhere (native losses nowhere exceed 4×10^{-3}): the ranking of instruments depends on whose generation the

planner weighs.

Who should conversion cover? The conversion channel as actually implemented in China is *skill-targeted*: points-based hukou systems and college-admission quotas grant residence to the educated. Appendix Table A2 evaluates this design — conversion made certain above a skill threshold (the top decile of adults, or the top half), the benchmark schedule kept below — and it fails on both of the margins that make conversion valuable. It captures almost none of the welfare: the top-decile policy delivers 4% of the universal gain, the median policy 42% (net of the tax cost), because the value of guaranteed access is *decreasing in the recipient's ability to self-insure* — skilled parents can substitute private inputs against enrollment risk, so selecting recipients on skill selects against value. And it reverses conversion's mobility effect: guaranteeing urban schooling to the children of skilled parents while leaving everyone else in the lottery *raises* the IGE, from 0.555 to 0.610 at the median threshold, where universal conversion lowers it to 0.381 — the same instrument flips sign on mobility with its coverage. Targeting's one virtue is a gentle transition: the top-decile policy is nearly painless for the cohorts alive at the reform. But it buys that gentleness at a worse welfare-per-dilution ratio than simply scaling universal conversion down, so if a gradual path is the goal, the margin to move is coverage breadth, not a skill threshold.

When conversion travels down the lineage. The benchmark holds birth city fixed across generations: a converted family is local within its lifetime, but its child enters adult life at the ancestral home. The data reject the fixed lineage: in the CFPS, children of converted rural-born mothers hold urban hukou with probability 0.93 — 0.79 already at age three — against 0.04 when the mother has not converted (Appendix E.5). We therefore evaluate the variant in which a converted family's child inherits the destination as its own home city with the measured probability. Transmission makes the birth distribution endogenous and turns the model into one of *hukou urbanization*: the rural lineage has no inflow, so it drains at the rate at which rural-born families convert, and the variant's stationary state is the fully urbanized economy. Three results. First, *the pace*: at the benchmark, 3.4% of rural-lineage children graduate in converted families per generation — a half-life of twenty generations, roughly four centuries. The status quo does dissolve rural status, and with it the left-behind category, by conversion alone — on a time scale that makes the point moot for anyone now alive. It also means the fixed-lineage benchmark is an accurate approximation at the horizons the welfare analysis weighs (each successive generation enters the dynastic calculus at weight $\gamma\beta^{J_c} \approx 0.17$), while the variant governs the demographic long run. Second, *the terminal state*: rural-born newborns vanish, lineages spread across the urban tiers, the left-behind share falls to 8.5% (residual separation among urban-origin families), and mean log graduation human capital exceeds the benchmark by half a log point — composition eventually does what policy otherwise would. In that world the access reform's

welfare gain shrinks to a 6% consumption equivalent from 29% (both before financing; ΔW +0.20 from +0.86): equal access is valuable exactly in proportion to the population that has not yet converted. Third, *access policy is also demographic policy*: equal access nearly doubles the rural lineage’s conversion outflow (3.4% $\rightarrow$ 6.3% per generation), halving the urbanization half-life from twenty generations to eleven, because removing enrollment risk pulls families into the cities where conversion happens. Access reform and residence conversion are thus *dynamic substitutes*: a planner can wait centuries for conversion to dissolve rural status on its own, or halve the wait while delivering the schooling gains to the generations of rural-born children who would otherwise grow up during it.

7 Winners and losers from education access restriction

The national case is settled by the preceding sections: equal access raises entrant welfare by 28% of permanent consumption net of the tax cost ($\Delta W = +0.84$; Table 12), and in the long run, at the estimated fiscal elasticity, every lineage shares in the gain. Nor is the gain an artifact of comparing two points: varying any single tier’s access over the full $[0, 1]$ range, with the rest of the benchmark schedule held fixed, national welfare rises monotonically everywhere (Figure 4, solid lines). This section asks the converse: *whom does restriction still serve?* We answer with an *incidence* exercise, not a political model: the model contains no political process, housing channel, or local-government objective, but it can measure the stake each birth lineage holds in each tier’s access policy — the input any political-economy account of the observed schedule would need. The text focuses on the megacities, where the observed restriction is harshest; Appendix G reports the scan for every tier.

The megacity case. Figure 4, panel (b), varies megacity access alone over $[0, 1]$. The megacity-born lineage’s welfare is *flat* across its entire own access range: it varies by less than ± 0.01 ($\pm 0.3\%$ of consumption) between full closure and full openness (-0.002 to $+0.031$ on the refined state space of Appendix I). By contrast, a real gain from closure exists in the small tiers (+12% of consumption for the tier-2-born, a lineage holding 0.6% of births); Appendix G reports that scan tier by tier. Under the estimated fiscal rule, what the lineage would gain in per-child spending by excluding migrant children is almost exactly offset by what it loses when restriction turns away the migrants whose taxes and human capital feed its budgets and its wage premium; only under rigid budgets ($\beta_g = 0$) does the megacity-born lineage again gain from closure. Everyone else’s stake in megacity access, meanwhile, is the largest in the schedule. Moving ρ_7 alone from its measured 0.57 to one delivers +14% of consumption nationally net of its tax cost ($\Delta W = +0.45$, over half of the entire equal-access gain) and a +23% gain (+0.69) to the rural-born lineage. Full megacity closure would cost the nation 16% of consumption ($\Delta W = -0.57$), and it would cost the *other urban* lineages heavily as well (tier-2-born -14% ,

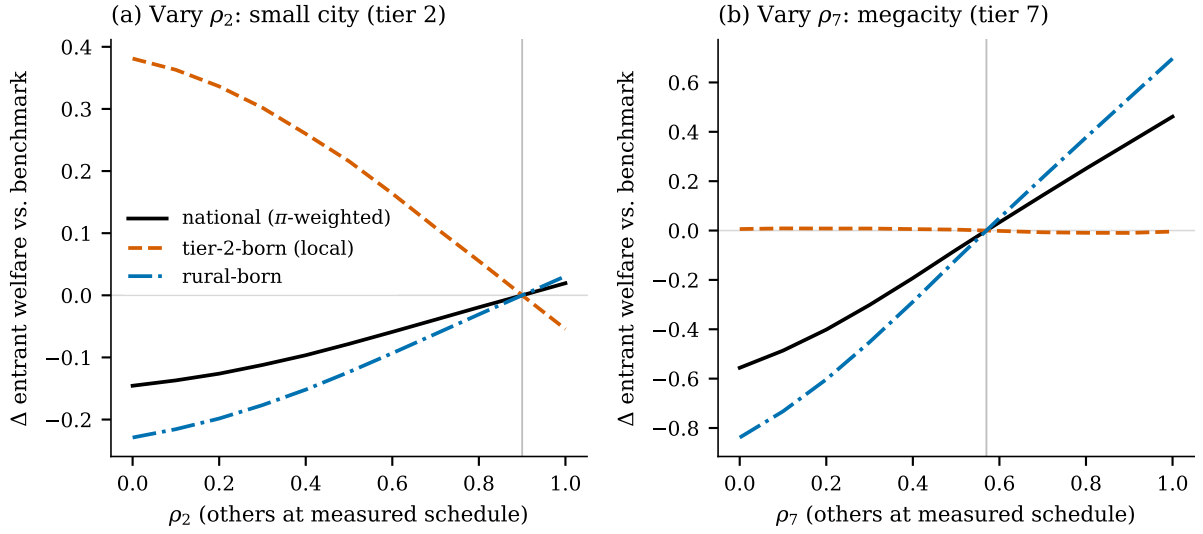


Figure 4: Local versus national stakes in one tier's access

Notes: Each panel varies one tier's access over $[0, 1]$, holding every other tier at its benchmark value, and plots the change in entrant welfare relative to the benchmark for the nation (population-weighted), the tier's own lineage, and the rural-born lineage. The vertical line marks the tier's benchmark ρ_i . Steady states under the estimated fiscal rule; the tax charge of a single-tier experiment (at most 0.015) is common to all lineages, so it shifts every curve equally and leaves the incidence comparisons unchanged.

tier-3-born -12% ; ΔW -0.49 and -0.44 : their descendants are among the migrants a closed megacity turns away).

Reading the observed schedule. In the data, restriction is mildest in the small tiers ($\rho_2 = 0.90$) and harshest in the megacities ($\rho_7 = 0.57$) — with the caveat that the mild end of the gradient is an assigned value, interpolated from the observed city shares and the national average, while the megacity value is computed from published city shares (Appendix F). The incidence map of Figure 4 is the mirror image. A lineage that benefits from restriction exists only where policy restricts *least*: small-tier incumbents gain from closing their schools to migrant children ($+12\%$ of consumption; Appendix G). Where policy restricts *hardest*, no lineage benefits at all: the megacity-born are indifferent to their own tier's access, while the nation's and the rural majority's stakes in megacity openness are the largest in the schedule. This inversion delivers the section's conclusion, which is a negative one. The standard justification for megacity enrollment barriers is that they protect incumbent children's school resources. The model can price that protection. At the fiscal-adjustment speed the data display, it is worth approximately nothing to the incumbents it is meant to protect: exclusion preserves some per-child spending, but it also turns away families whose taxes and human capital would have fed the same budgets, and the two effects cancel. Whatever does sustain megacity barriers must therefore operate through channels outside this model — housing and congestion pressures, service budgets beyond education, or the interests of current residents rather than birth lineages. What the model can say is

that the education-budget rationale is not it.

8 Conclusion

We have studied the welfare implications of migrant children's access to urban public schooling in a spatial, dynastic economy with endogenous human capital, agglomeration, and a schooling externality. Access is modeled as the institution operates: an enrollment probability drawn after families choose where to live, with school budgets adjusting to enrollment at the pace we estimate in three decades of city data. The lottery formulation is not cosmetic. A certainty-equivalent discount produces an economy with no excluded children — the state that holds a quarter of brought migrant children in the data — and delivers only two thirds of the family-arrangement response to access policy and four fifths of the reform's welfare gain.

Removing enrollment barriers is a family-reunification and mobility reform whose burden falls on a generation, not on a class. Migrant families gain in every urban tier; natives of the restrictive tiers bear a transitory budget dip (-2.4% of consumption for megacity natives alive at the reform); and in the long run budgets recover, families reunify, mobility rises, and every birth lineage is better off at the estimated fiscal elasticity — by 28% of permanent consumption for the average dynasty — though at the conservative end of the estimated range the most protected lineage is approximately unharmed rather than a net gainer. Financing governs who pays: under rigid budgets the elite-tier lineages would lose permanently, while funding that follows each child shrinks reform-date native losses to a fraction of a percent of consumption. The incidence of restriction itself is sharply asymmetric: a lineage that gains from its own restriction exists only in the small tiers, while in the megacities — where restriction is harshest — the protected lineage is indifferent to its own protection and the nation's stake in openness is the largest in the schedule. And on the longest horizon, the status quo dissolves rural status by conversion alone only on a four-century clock; equal access halves that wait, while delivering the schooling gains to the generations of rural-born children who would otherwise grow up during it.

None of these lessons is specific to the institution that makes them measurable: wherever school finance is local and enrollment follows residence, the same margins — enrollment risk, family separation, fiscal adjustment, and the skilled family's ability to self-insure — decide who gets schooled, with whom, and at whose expense. Two extensions are natural. An explicit account of current residents' interests — housing, property values, and non-school congestion — would let the incidence analysis speak to the politics of the observed schedule, and richer heterogeneity in the education technology would let the model speak to the composition, not only the level, of the human-capital gains. We leave these to future work.

References

- All-China Women's Federation (2013). Research report on rural left-behind children and urban–rural migrant children in China. Technical report, All-China Women's Federation. In Chinese.
- Anstreicher, G. (2024). Spatial influences in upward mobility. *Journal of Political Economy*, 132(1):149–199.
- Antman, F. M. (2013). The impact of migration on family left behind. In *International Handbook on the Economics of Migration*, pages 293–308. Edward Elgar.
- Bénabou, R. (1996). Equity and efficiency in human capital investment: The local connection. *Review of Economic Studies*, 63(2):237–264.
- Browning, M., Hansen, L. P., and Heckman, J. J. (1999). Micro data and general equilibrium models. In *Handbook of Macroeconomics*, volume 1A, pages 543–633. Elsevier.
- Chen, Y., Feng, S., and Han, Y. (2020). Choices or constraints: Education of migrant children in urban China. *Population Research and Policy Review*, 39(4):671–690.
- Combes, P.-P. and Gobillon, L. (2015). The empirics of agglomeration economies. In *Handbook of Regional and Urban Economics*, volume 5, pages 247–348. Elsevier.
- de la Roca, J. and Puga, D. (2017). Learning by working in big cities. *Review of Economic Studies*, 84(1):106–142.
- Durlauf, S. N. (1996). A theory of persistent income inequality. *Journal of Economic Growth*, 1(1):75–93.
- Eckert, F. and Kleineberg, T. (2021). Saving the American Dream? education policies in spatial general equilibrium. Policy Research Working Paper 9574, World Bank.
- Epple, D. and Romano, R. E. (1998). Competition between private and public schools, vouchers, and peer-group effects. *American Economic Review*, 88(1):33–62.
- Fan, Y., Yi, J., and Zhang, J. (2021). Rising intergenerational income persistence in China. *American Economic Journal: Economic Policy*, 13(1):202–230.
- Fernández, R. and Rogerson, R. (1996). Income distribution, communities, and the quality of public education. *Quarterly Journal of Economics*, 111(1):135–164.
- Fernández, R. and Rogerson, R. (1998). Public education and income distribution: A dynamic quantitative evaluation of education-finance reform. *American Economic Review*, 88(4):813–833.

- Guo, N., Huang, Z., Lu, M., Xi, X., and Zhong, Y. (2026). Education policy, family migration, and the service sector in China. Work in progress; slides, Xiamen University, May 2026.
- Huang, Z. and Zhang, J. (2021). School restrictions, migration and peer effects: A spatial equilibrium analysis of children's human capital in China. Working paper, SSRN 3678300.
- Imbert, C., Monras, J., Seror, M., and Zylberberg, Y. (2023). Floating population: Migration with(out) family and the spatial distribution of economic activity. Working Paper 2023-26, Federal Reserve Bank of San Francisco.
- Lee, S. Y. and Seshadri, A. (2019). On the intergenerational transmission of economic status. *Journal of Political Economy*, 127(2):855–921.
- Lu, Y. and Zhou, H. (2013). Academic achievement and loneliness of migrant children in China: School segregation and segmented assimilation. *Comparative Education Review*, 57(1):85–116.
- Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22(1):3–42.
- Ma, L. and Tang, Y. (2024). The distributional impacts of transportation networks in China. *Journal of International Economics*, 150.
- Meng, X. and Yamauchi, C. (2017). Children of migrants: The cumulative impact of parental migration on children's education and health outcomes in China. *Demography*, 54(5):1677–1714.
- Ngai, L. R., Pissarides, C. A., and Wang, J. (2019). China's mobility barriers and employment allocations. *Journal of the European Economic Association*, 17(5):1617–1653.
- Restuccia, D. and Urrutia, C. (2004). Intergenerational persistence of earnings: The role of early and college education. *American Economic Review*, 94(5):1354–1378.
- Sieg, H., Yoon, C., and Zhang, J. (2020). The impact of migration controls on urban fiscal policies and the intergenerational transmission of human capital in China. Working Paper 27764, National Bureau of Economic Research.
- Tombe, T. and Zhu, X. (2019). Trade, migration, and productivity: A quantitative analysis of China. *American Economic Review*, 109(5):1843–1872.
- Zhang, H., Behrman, J. R., Fan, C. S., Wei, X., and Zhang, J. (2014). Does parental absence reduce cognitive achievements? evidence from rural China. *Journal of Development Economics*, 111:181–195.

Online Appendix

Left Behind by Design: School Access and Family Migration

Yang Tang and Jingwei Zhou

This online appendix collects the notation, model variants and robustness checks, the empirical appendix, the access-schedule construction, the single-tier incidence scan, and computational details. Numbering of appendix sections, tables, and figures is prefixed by the section letter or “A”.

A Notation

Table A1: Model notation

<i>Individual states</i>	
j	age, $j = J_c + 1, \dots, J$ over the working life
i	current work city
i_b	birth / home city (fixed lineage attribute)
$i_c \in \{i_b, i\}$	child's current city
$a \in \{0, 1\}$	migrant child's enrollment status this period (schooled in home city: $a = 1$)
h	adult human capital (fixed after labor-market entry)
h_c	child's human capital (h_1 at birth)
y	beginning-of-period financial wealth
<i>Choices</i>	
c, b'	consumption; saving, $y' = (1 + r) b'$
m, t	education money input; parental time input
i'	next period's work city
$i'_c \in \{i_b, i'\}$	child's next city (leave at home vs. bring)
<i>Equilibrium objects</i>	
w_i	wage per efficiency unit of labor
N_i	aggregate efficiency labor
$\bar{H}_i$	city-average adult human capital (spillover)
g_i	per-enrolled-child public education spending, eq. (7)
$g(i_c, i_b)$	spending a child receives: g_{i_c} if $i_c = i_b$ or enrolled ($a = 1$); 0 if excluded
$W(i, h)$	lifetime (dynastic) value of a labor-market entrant
<i>Policy</i>	
s_E	education budget share (anchor of the allocation rule)
τ	labor-income tax rate balancing the national education budget, eq. (8)
$\rho_i \in [0, 1]$	migrant child's public-school enrollment probability in city i
$p(i, h) = \min\{1, \bar{p}_i h^\psi\}$	residence-conversion probability (h -increasing); re-registers the home city
$\theta_i \geq 1$	skill bias of urban efficiency units (income $w_i h^{\theta_i}$; $\theta_1 = 1$)
<i>Shocks and mobility costs</i>	
ξ	education-technology shock, $\ln \xi \sim N(0, \sigma^2)$, i.i.d.
z, ϕ, ω	education technology: scale, money share, curvature, eq. (5)
$\iota_{i'}$	Gumbel location-preference shock, scale κ
$\zeta_{\text{leave}}, \zeta_{\text{bring}}$	Gumbel bring-child shocks, scale κ_c
$d_{ii'}$	migration cost (utility units)
m^c	settlement cost of keeping the child in the city (utility units)
<i>Value functions</i>	
v_n	no-child period value
v_c	with-child period value
V_c	bring-child-smoothed continuation
i_b (re-registered)	set to the work city on conversion: lottery off, no settlement cost
$\mathcal{A}$	arrival value: $p V_c(\cdot, i_b=i') + (1 - p) V_c(\cdot, i_b)$, mixed over the conversion draw

Notes: The reduced-form sections (Section 2 and Appendix E) use conventional regression notation — $\beta, \gamma, \delta_t, \vartheta_{g,s}, \alpha_c, \varphi$ — whose letters occasionally coincide with model parameters; within those sections the regression reading applies throughout.

B Skill-targeted residence conversion

Table A2 reports the counterfactuals behind the coverage discussion of Section 6: conversion made certain for parents above a skill threshold h^* — $p(i, h) = 1$ for $h \geq h^*$ wherever conversion operates — with the benchmark schedule $\bar{p}_i(h/h_{\text{ref}})^\psi$ retained below the threshold.

Table A2: Skill-targeted conversion: certainty above the threshold h^* , benchmark schedule below

	benchmark	top decile $h^* = 3.0$	top half $h^* = 1.5$	universal $p = 1$
Co-residence rate, rural migrants	0.375	0.391	0.552	0.806
Left-behind share (all children)	0.187	0.180	0.123	0.059
Mean log child graduation HC	0.010	0.030	0.219	0.511
IGE (child–parent human capital)	0.555	0.569	0.610	0.381
Urban share, top-skill-third rural-born	0.496	0.499	0.522	0.546
Tax rate τ , long run (%; anchor 8.0)	8.0	8.1	8.7	9.7
Entrant welfare, π_{birth} -weighted	−0.138	−0.054	0.704	1.871
Δ vs. benchmark, net of the tax cost (CEV)	—	+0.08 (+2.5%)	+0.82 (+28%)	+1.95 (+79%)
rural-born (W_1)	−3.549	−3.449	−2.438	−0.660
megacity-born (W_7)	7.287	7.253	7.473	7.919
Megacity spending trough, transition (%)	—	−2.6	−13.2	−27.2
ΔV , megacity natives at reform (10^{-3})	—	−23	−80	−136

Notes: Steady states under the estimated fiscal rule ($\beta_g = 0.51$). Thresholds are grid-exact: $h^* = 3.0$ grants certainty to the top $\approx 11\%$ of adults by mass, $h^* = 1.5$ to the top $\approx 48\%$. Steady-state rows as in Table 12; the last two rows are transition objects from the corresponding perfect-foresight runs (the trough of megacity per-enrolled-child spending, and the reform-date mean value change of megacity native families as in Table 9).

C Private instruction and the excluded state

In the benchmark technology (5), household money enters only the goods component: an excluded migrant child is educated with purchased materials and parental teaching, and a left-behind child with materials and the home public input alone. Money cannot hire a tutor. This is a strong restriction — it makes parental human capital the only source of private instruction, which in principle could manufacture the self-insurance gradient the mechanism relies on. This appendix relaxes it: household money may purchase instructional labor at price χw_i per efficiency unit — the same units in which public instruction enters — in every child state, including excluded children (private tutoring or migrant-school instruction) and left-behind children (hired tutoring at home). The household buys instruction whenever it is cheaper than parental time or the time endowment binds; $\chi \rightarrow \infty$ recovers the benchmark.

Two findings. First, *the data reject cheap private instruction*. At $\chi = 1$ the per-draw enrollment odds ρ_i are unchanged, yet the excluded share collapses to 5% — an equilibrium effect running through conversion: tutoring-fueled human capital pushes migrant parents over

Table A3: Purchasable instruction: the tutoring premium and the benchmark fit

	benchmark	$\chi = 1$	$\chi = 2$	$\chi = 3$	$\chi = 3, m^c$ re-fit	target
Excluded share, brought migrants	0.24	0.05	0.16	0.21	0.22	0.2–0.3
Co-residence rate, rural migrants	0.375	0.52	0.31	0.27	0.375	0.37
Rural→urban migration rate	0.431	0.62	0.50	0.44	0.47	0.426
SD(ln labor income)	1.33	1.48	1.39	1.32	1.27	1.25
IGE (child–parent human capital)	0.555	0.77	0.70	0.65	0.60	0.4–0.6
Education expenditure share	0.120	0.134	0.134	0.134	0.131	0.120
Mean log child graduation HC	0.01	1.73	0.87	0.42	0.32	

Notes: Stationary equilibria at unchanged benchmark parameters except as indicated; the last column re-fits the settlement cost m^c ($1.46 \rightarrow 1.08$) to restore the co-residence target at $\chi = 3$. At $\chi = 1$ (tutoring at the market wage) the economy is transformed — exclusion nearly costless, migration 19 points higher, child human capital more than doubled, intergenerational persistence far above its empirical range: the calibration targets reject cheap purchasable instruction. The fit approaches the data only when private instruction carries a premium of roughly three times the market wage — consistent with tutoring being retail-priced skilled labor against the public school’s economies of scale.

the education-scored conversion schedule (the rural lineage’s conversion flow rises from 3.4% to 46% per generation), and converted families’ children enroll as natives outside the lottery. Migration and co-residence surge past their targets, and measured persistence and dispersion leave their empirical ranges (Table A3): an economy in which families could hire instruction at the market wage would not look like China’s. Matching the calibration targets requires a tutoring premium of roughly $\chi \approx 3$. Notably, what expensive-but-available instruction changes most is not the enrollment margin — the excluded share is back near benchmark by $\chi = 3$ — but the *left-behind* margin: home tutoring substitutes for parental co-residence, which is why the settlement cost re-fits downward ($m^c = 1.08$ against 1.46). Second, *the paper’s conclusions survive at the data-consistent premium*. With $\chi = 3$ and the co-residence target restored, the equal-access reform under the estimated fiscal rule delivers a welfare gain of +0.78, against +0.86 in the benchmark (both before tax charges, which are common to the comparison). Every lineage again gains (megacity-born +0.18, rural-born +1.13); co-residence rises 6.0 points; the IGE falls by 0.11. And the model’s response to the 2014-style tightening episode is +3.9 percentage points against the benchmark’s +4.1 — the policy elasticity validated against the census event study is essentially invariant to purchasable instruction.

D The mechanism at a glance

Figure A1 collects the within-period sequence and the equilibrium feedback in one diagram.

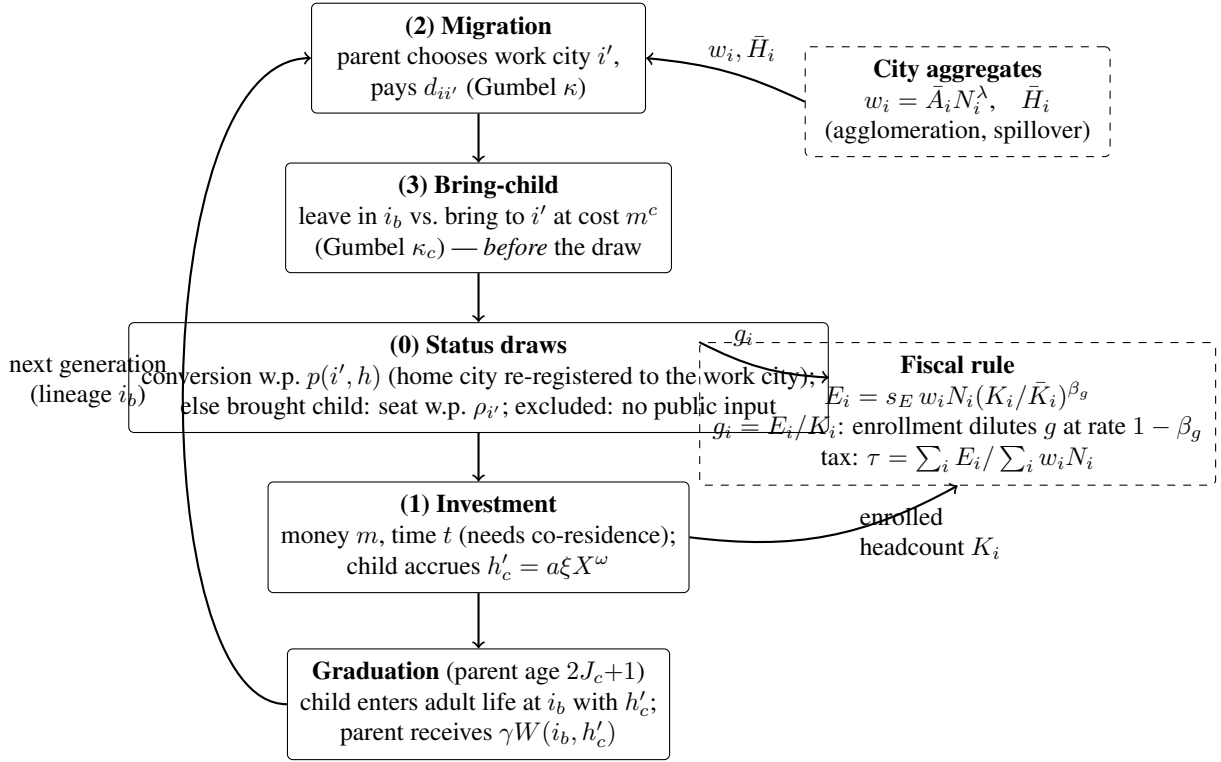


Figure A1: The mechanism in one page

Notes: Solid boxes: the family's within-period sequence (numbered as in Section 3); the bring decision precedes the enrollment draw, which is what makes access a risk rather than a discount. Dashed boxes: equilibrium objects. The feedback loop on the right is the dilution channel: enrollment expands K_i , budgets respond with elasticity β_g , and per-child spending g_i moves the value of the next family's draw.

E Empirical appendix

This appendix reports the full regressions and sample constructions behind the motivating evidence of Section 2 and the measurement used in the calibration (Section 4). All scripts and outputs are archived with the replication package.

E.1 Family separation and destination restrictiveness (Fact 2)

Sample. 2005 one-percent population census microdata. Mothers aged 20–39, currently married, agricultural hukou, registered outside their county of residence, residing in a city or town, household head or spouse, with at least one surviving child; $n = 18,214$ across 283 destination cities. A child is left behind if the mother's surviving children exceed her co-resident children. Destination restrictiveness groups follow a standard city-tier classification; estimation uses census sampling weights, with standard errors clustered by destination city. Table A4 reports the estimates of equation (1) across specifications.

Table A4: Fact 2 robustness: left-behind probability, destination restrictiveness, and mother’s education

Specification	R_c	$R_c \times \text{LowEd}_m$	n
(1) Baseline (eq. 1)	2.3 (2.2)	4.9 (1.8)	18,214
(2) + migration distance (inter-prov., duration)	2.9 (2.2)	4.5 (1.8)	18,214
(3) + spouse presence and education	2.2 (2.0)	4.7 (1.7)	18,214
(4) Fractional outcome (share left behind)	2.6 (2.1)	4.9 (1.7)	18,214
(5a) Mothers 28–39 (school-age children)	2.0 (2.5)	5.5 (2.0)	13,960
(5b) Mothers 20–27 (mostly preschool)	3.3 (3.1)	2.7 (3.4)	4,254
(6) Continuous log city size in place of R_c	−3.0 (2.9)	5.5 (1.8)	18,187

Notes: Coefficients $\times 100$ (percentage points per restrictiveness-group step; row 6: per log point of destination city size). Standard errors in parentheses, clustered by destination city; census weights. All specifications control for a quadratic in the mother’s age and the LowEd main effect. A probit version of the baseline yields average marginal cross-differences of +11 to +15 points from the least to the most restrictive group.

E.2 School budgets and enrollment (Fact 3)

Sample. Prefecture-level city panel, 1990–2023 vintage (estimation window 1993–2022 after differencing and trimming), 299 cities, assembled from the standard city statistical-yearbook compilation: education expenditure, total local fiscal expenditure, primary and regular-secondary enrollment, regular-college enrollment, and full-time primary-plus-secondary teachers. The full regressions appear in Table 2 in the main text. Complementing the panel, a 2005 cross-section of per-student spending on the city’s migrant share (inter-county migrants from the 2005 mini-census, controlling log fiscal expenditure per registered resident, with province fixed effects) shows the same partial adjustment in levels: cities with larger migrant inflows spend less per enrolled student.

E.3 Child cognitive outcomes by schooling state (CEPS)

Sample. China Education Panel Survey, baseline 2013–14, $n = 16,284$ seventh- and ninth-graders with the standardized 3PL cognitive score; states defined by hukou, migration, and school type (public versus non-public). Regressions condition on gender, grade, and both parents’ education; the omitted state is rural children living with both parents. Table A5 reports levels and one-year value-added.

E.4 Residence conversion and education (CFPS)

Sample. CFPS adults born with agricultural hukou (measured at age 3), stacked waves 2010–2018, $n = 18,910$. Table A6 reports the share holding non-agricultural hukou at interview, by educational attainment, and probit average marginal effects controlling for birth cohort, gender, and childhood province.

Table A5: CEPS cognitive scores by child state

	Baseline level	Value-added (wave 2 wave 1)
Urban native	+0.12 (0.03)	−0.02 (0.03)
Rural, left behind	+0.01 (0.02)	−0.00 (0.04)
Migrant, public school	+0.08 (0.04)	−0.02 (0.08)
Migrant, non-public school	−0.15 (0.15)	+0.07 (0.13)

Notes: Reference group: rural children living with both parents. Standardized 3PL cognitive score; controls for gender, grade, and parents' education; school-clustered standard errors. The migrant non-public cell is thin (228 students in seven schools), reflecting the survey frame's limited coverage of migrant schools; the level gaps, not the one-year value-added, are the object the model's states speak to (Section 4.3). The enrolled–excluded contrast quoted in the text (−0.20, s.e. 0.07) is the within-migrant difference between the last two rows estimated directly.

Table A6: Hukou conversion by education, rural-born adults (CFPS)

Education	Converted share	<i>n</i>
Illiterate/semi-literate	0.05	5,570
Primary school	0.08	4,660
Junior high school	0.16	5,761
Senior high school	0.39	2,028
2/3-year college	0.71	566
4-year college	0.87	303

Notes: Probit average marginal effects relative to primary or less, with birth-cohort, gender, and childhood-province controls: junior high +12 points, senior high +31, college +73. Among adults still holding agricultural hukou at age 12, subsequent conversion rises from 6% to 75% across the same education groups.

E.5 Intergenerational hukou transmission

The lineage-transmission variant of Section 6 is disciplined by a direct CFPS measurement: children aged 0–15 linked to their parents through the family roster, with conversion defined as in Appendix E (agricultural hukou at age three, non-agricultural at interview). Among children of rural-born mothers, the probability of holding non-agricultural hukou is 0.93 when the mother has converted ($n = 713$) — and 0.79 measured already at the child’s age three — against 0.04 when she has not ($n = 5,267$). The father-side rates are 0.81 and 0.03; conditioning on both parents rural-born, a family with at least one converted parent transmits urban status with probability 0.76, against 0.01 for families with none. Conversion, once obtained, passes to the next generation nearly one-for-one; the variant sets the transmission probability to the mother-based 0.93.

E.6 Migration selection and returns to schooling (census and CFPS)

The migration-selection gradient of Section 4.3 — the urban-residence share of rural-born adults by education, 0.30/0.48/0.65/0.77 from primary-or-less to college ($n = 12,007$) — is estimated in the CFPS on rural-born adults regardless of subsequent conversion, with probit AMEs of +20, +40, and +54 points relative to primary or less under age, gender, and childhood-province controls. The returns-to-schooling anchor for θ_u is estimated in the 2005 census (log monthly income on years of schooling, quadratic in age, gender; census weights, prefecture-clustered standard errors; employed adults 16–60). Among agricultural-hukou workers the return is 6.3% per year of schooling in cities (s.e. 0.6, $n = 164,804$) against 4.7% in the countryside (s.e. 0.3, $n = 676,132$) — the composition-held ratio of 1.34 used in the calibration. The unrestricted all-worker returns are 10.7% and 5.6% (ratio 1.9), inflated by compositional selection into cities.

E.7 The 2014 tightening: reform-based event studies

Design. The 2014 National New Urbanization Plan assigned residence policy by city size: strict points-based admission in the largest cities, mandated relaxation below. We pool the 2005 and 2015 mini-censuses (1% intercensal surveys) with a two-million-person extract of the 2010 census and construct, in each wave, the migrant-mother sample of Section 2: rural mothers aged 20–39, currently married, registered outside their county of residence, household head or spouse, at least one surviving child; a child is left behind if surviving children exceed co-resident children counted from the household roster. Treatment groups follow the destination restrictiveness classification of Table 1: T1 (Beijing/Shanghai/Guangzhou/Shenzhen, strict points systems), T2 (major second-tier cities, partially treated), control (other and small prefectures, mandated relaxation). The event-study specification regresses the left-behind indicator on group $\times$ wave interactions with 2005 as base, wave and group main effects, a quadratic

in age, and the low-education indicator, clustered by destination prefecture. The three waves bracket two opposite-signed events: the 2005–2010 leg spans the 2008 borrowing-fee abolition and the 2010–2015 leg the 2014 tightening, so both wave interactions are treated-leg coefficients rather than a placebo and an effect. Definitional notes: the 2015 questionnaire dropped the hukou-nature question (rural status in 2015 is rural land-contract rights, M45); 2005’s city/town-residence filter has no counterpart in the other waves and is dropped throughout (the 2005 estimates are robust to it); estimates are unweighted (the 2010 extract carries no weights).

Table A7: Family separation around the 2014 tiered tightening: raw cell means

Destination group	Less-educated mothers				More-educated mothers			
	2005	2010	2015	Δ_{10-15}	2005	2010	2015	Δ_{10-15}
First tier (strict points systems)	0.71	0.84	0.65	–20	0.58	0.76	0.43	–33
Major second tier (partial)	0.58	0.78	0.45	–34	0.45	0.70	0.28	–42
Other prefectures (relax)	0.63	0.79	0.45	–33	0.57	0.70	0.28	–42
Smallest prefectures (relax)	0.41	0.68	0.14	–53	0.45	0.62	0.12	–50

Notes: P(any child left behind) among migrant mothers, by destination restrictiveness group, census wave, and mother’s education (less-educated: junior high or below); Δ_{10-15} in percentage points. Sample construction as in the text; unweighted cell means. Cell sizes range from 168 to 7,996 (1,165–3,149 in 2015 for the less-educated panel). Among the less-educated — the group the model’s self-insurance mechanism centers on — the first tier’s 2010–2015 decline is smallest by a wide margin and the smallest prefectures’ largest, with the two middle groups within a point of each other; the more-educated panel shows the same first-tier-versus-relaxation contrast.

Table A8: Family separation around the 2014 tiered tightening: event-study estimates

	× 2010 (liberalization leg)	× 2015 (tightening leg)
First tier (strict points systems)	–5.2 (2.3)	+11.8 (2.6)
Second tier (partially treated)	+1.7 (3.3)	+6.1 (3.6)
First tier × low-education (triple)	–10.7 (2.6)	–2.2 (3.2)
Second tier × low-education (triple)	–9.9 (3.9)	–7.0 (3.9)

Notes: Coefficients ×100 (percentage points), relative to the small-city control group and the 2005 wave; standard errors in parentheses, clustered by destination prefecture. $n = 54,996$ mothers across the three waves (22,052 / 23,109 / 9,835). Top panel: the event study described in the text. Bottom panel: triple-difference interactions with the mother’s low-education indicator — the post-2014 tightening does not load differentially on less-educated mothers, unlike the 2005 cross-section interaction of Table 1. Table A7 reports the underlying cell means. The share of migrant mothers observed in first-tier destinations falls from 0.33 (2005) to 0.19 (2015): the extensive margin responds alongside family arrangements, so remaining-migrant composition is not fixed.

Inference with four treated cities, and the 2015 definition. Two features of the design deserve explicit treatment. First, treatment varies across four destination groups — the strictly treated group is four cities — so prefecture-clustered standard errors, while conventional, do not account for shocks common to a treatment group. Table A9 reports two complements. A randomization test permutes the tier-group labels across the 287 destination prefectures (holding

group counts fixed, 999 draws) and re-estimates the event study: the observed tightening coefficient has a permutation p -value of 0.23, and the sign-reversal statistic $\beta_{T1,2015} - \beta_{T1,2010} = +17.0$ has $p = 0.22$ — placebo groups of four random prefectures often produce large swings, because small destinations’ separation rates move by tens of points across waves. The permutation exercise is conservative for this design in one specific sense: actual assignment was mechanical in city size, and the causal argument rests not on size-random assignment but on the response flipping sign with the direction of policy on the policy dates. The continuous-dose specification identifies that reversal from all 287 prefectures rather than four groups: interacting the destination’s log 2010 registered population with the waves yields -1.5 (s.e. 1.9) per log point $\times$ 2010 and $+3.5$ (s.e. 2.0) $\times$ 2015 — the same sign flip, identified off the full size gradient. At the first tier’s mean log-size gap over the control group (0.93 log points), the implied tightening effect is $+3.3$ points (s.e. 1.9) — smaller than the group estimate, consistent with points systems binding disproportionately in the very largest cities beyond a log-linear dose.

Second, the 2015 questionnaire measures rural status by land-contract rights rather than hukou nature, so the tightening leg’s sample definition changed in the wave that carries it. Two bounds. Reweighting the 2015 sample to the 2010 joint distribution of destination group and mother’s education moves the tightening coefficient from $+11.8$ to $+12.7$ (s.e. 2.7). Dropping the rural filter entirely — all inter-county migrant mothers, a sample definition untouched by the questionnaire change — gives $+6.1$ (s.e. 2.4), the attenuation expected from adding urban-hukou mothers whom the restriction does not bind, with the sign reversal intact (-4.0 in the liberalization leg, $+6.1$ in the tightening leg). First-tier rural-sample cell sizes are 7,347 (2005), 4,702 (2010), and 1,819 (2015).

The school-system counterpart. In the city panel of Table 2 we estimate equation (3) for the same treatment groups; Figure A2 plots the full coefficient paths for both events, and Table A11 reports selected coefficients. Panel (a): after the 2008 abolition of migrant “borrowing fees” in public schools, large-city enrollment rises steadily relative to small cities — $+15$ log points in the second tier by 2013 (s.e. 2.9) on flat pre-trends, with the same but noisier path in the first tier. Panel (b): the first tier’s relative rise — roughly 4–5 log points per year through 2012 — breaks after 2014. Table A12 estimates the break rather than reading it off the figure. Against each group’s own linear pre-trend fitted on 2008–2013 and extrapolated forward, the first tier’s path falls progressively below trend: -17 log points by 2017 and -40 by 2022, individually significant from 2017 on, with a mean post-period deviation of -19.5 (s.e. 7.4). The partially treated second tier stays on its trend (mean deviation 0.0, s.e. 3.5). The first-minus-second-tier contrast, which nets out any trend common to large cities, is -7.2 (s.e. 2.9) by 2016 and -11.3 (s.e. 4.8) by 2018.

Table A9: Group-level inference and the 2015 sample definition

	estimate (s.e.)	RI p
<i>Randomization inference (999 permutations of tier labels)</i>		
first tier $\times$ 2015	+11.8 (2.6)	0.23
first tier $\times$ 2010	−5.2 (2.3)	0.66
sign reversal: T1 $\times$ 2015 − T1 $\times$ 2010	+17.0	0.22
<i>Continuous dose: log 2010 city population $\times$ wave (287 prefectures)</i>		
$\times$ 2010	−1.5 (1.9)	
$\times$ 2015	+3.5 (2.0)	
implied first-tier 2015 effect (log-size gap 0.93)	+3.3 (1.9)	
<i>The 2015 leg under alternative sample definitions (T1$\times$2015)</i>		
baseline (rural by land-contract rights)	+11.8 (2.6)	
2015 reweighted to 2010 group $\times$ education	+12.7 (2.7)	
no rural filter, all waves ($n = 79,231$)	+6.1 (2.4)	

Notes: Percentage points. Randomization inference permutes the tier-group labels across the 287 destination prefectures, holding the number of prefectures per group fixed, and re-estimates equation (2) on each draw; p is the two-sided share of draws with a larger absolute statistic. Continuous dose replaces the group dummies with the destination’s log 2010 registered (hukou) population, centered, interacted with the waves; standard errors clustered by prefecture. The no-rural-filter row includes all inter-county migrant mothers regardless of hukou type in every wave — a definition unaffected by the 2015 questionnaire change — and attenuates mechanically by including mothers the restriction does not bind.

Table A10: School-age placebo: event-study effects by the age of the mother’s children

	$\times$ 2010 (2008 liberalization leg)		$\times$ 2015 (2014 tightening leg)	
	preschool	school-age	preschool	school-age
First tier	−2.7 (3.5)	−6.1 [−3.3 (2.3)]	+12.6 (3.6)	+11.3 [−1.3 (3.1)]
Second tier	+1.8 (3.5)	+1.7 [−0.1 (2.8)]	+8.6 (4.2)	+5.3 [−3.4 (3.3)]

Notes: Group $\times$ wave effects (percentage points $\times 100$) from equation (2) fully interacted with a school-age indicator: mothers aged 28–39, whose children are mostly of compulsory-school age, versus 20–27 (mostly preschool) — the same proxy as the age split of Table 1. School-age cells report the preschool effect plus the interaction, with the interaction coefficient and its clustered standard error in brackets. The 2008 liberalization’s response doubles for school-age mothers; the 2014 tightening loads on both groups, consistent with points systems requiring documentation histories accumulated before school entry.

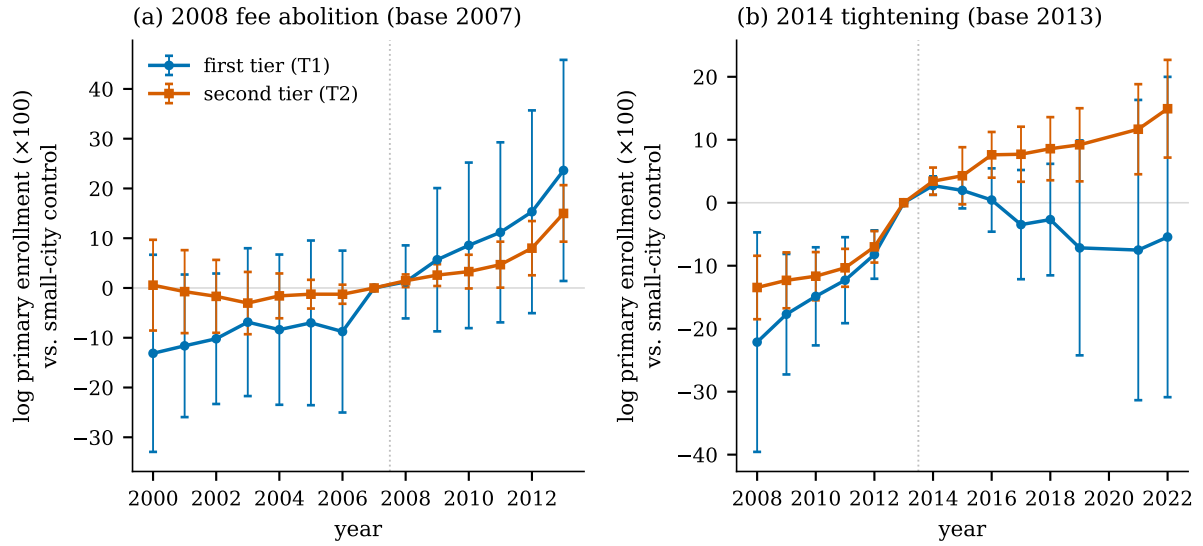


Figure A2: Enrollment event studies (equation (3)): log primary enrollment of the first-tier and major-second-tier city groups relative to the small-city control, around the 2008 borrowing-fee abolition (panel a, base 2007) and the 2014 tiered tightening (panel b, base 2013). City and year fixed effects and a log registered-population control; bars are 95% confidence intervals, clustered by city; 2020 and 2023, for which the yearbook enrollment counts are unavailable, are omitted from the plotted series and the connecting line.

Table A11: Enrollment event studies: selected coefficients (equation (3))

2008 fee abolition (base 2007)			2014 tightening (base 2013)		
Year	First tier	Second tier	Year	First tier	Second tier
2000	−13.1 (10.1)	+0.6 (4.7)	2008	−22.1 (8.9)	−13.5 (2.6)
2003	−6.9 (7.6)	−3.0 (3.2)	2010	−14.9 (4.0)	−11.7 (2.0)
2006	−8.8 (8.3)	−1.3 (1.0)	2012	−8.2 (2.0)	−7.0 (1.3)
2008	+1.2 (3.7)	+1.5 (0.6)	2014	+2.7 (0.8)	+3.4 (1.1)
2010	+8.6 (8.5)	+3.3 (1.7)	2016	+0.4 (2.6)	+7.6 (1.9)
2013	+23.6 (11.3)	+15.0 (2.9)	2018	−2.7 (4.5)	+8.6 (2.6)
			2022	−5.4 (13.0)	+14.9 (4.0)

Notes: Coefficients $\vartheta_{g,s} \times 100$ (log points) with standard errors in parentheses, clustered by city; selected years from the full paths of Figure A2. Each panel is estimated on its own window ($n = 3,952$ and $4,001$ city-years, 287 cities). The +15.0 (2.9) entry — second tier, 2013, left panel — is the second-tier enrollment gain relative to small cities cited in Section 2.

Table A12: The 2014 trend break, estimated: post-2014 enrollment relative to extrapolated group pre-trends

	First tier (dev. from own pre-trend)	Second tier (dev. from own pre-trend)	First — second tier (raw coefficients)
2014	+1.1 (2.5)	+4.1 (1.5)	−0.7 (1.2)
2015	−3.7 (4.5)	+2.5 (3.0)	−2.3 (2.6)
2016	−9.3 (6.1)	+3.5 (2.7)	−7.2 (2.9)
2017	−17.2 (7.4)	+1.1 (3.2)	−11.2 (4.7)
2018	−20.5 (8.6)	−0.4 (3.7)	−11.3 (4.8)
2019	−29.0 (9.5)	−2.2 (4.4)	−16.4 (9.0)
2021	−37.5 (11.0)	−4.6 (5.4)	−19.2 (12.5)
2022	−39.5 (11.7)	−3.7 (6.0)	−20.4 (13.3)
mean post-2013	−19.5 (7.4)	+0.0 (3.5)	—

Notes: Log points $\times 100$, delta-method standard errors from the full clustered coefficient covariance in parentheses. Columns 1–2: each group’s linear pre-trend is fitted on its 2008–2013 event-study coefficients (slopes +4.05 and +2.42 log points per year for the first and second tier) and extrapolated forward; entries are the post-2014 coefficients minus the extrapolation — the estimated counterpart of the “rise stops abruptly” reading of Figure A2. Column 3: the first-minus-second-tier difference in raw coefficients, which nets out any trend common to large cities. Specification and sample as in Table A11.

E.8 The fiscal elasticity: instrumented estimates

The elasticity β_g is identified in Fact 3 from within-city enrollment variation, which could respond to budgets, school quality, or local conditions. Table A13 re-estimates specification (4) by two-stage least squares under two instruments that carry only predetermined variation. The *demographic* instrument predicts a city’s school-age population from cohorts already counted in the prior census: children aged 6–15 in year t were born by the census date, so the city’s age pyramid, shifted forward, predicts enrollment mechanically (each census predicts the six years that follow it, so every school-age cohort is observed; city intercepts are census-window specific). The *policy* instrument interacts a city’s predetermined migrant exposure (the 2005 mini-census migrant share of adults) with the post-2014 tightening — an admission rule moves the education budget only through who is enrolled, which is the exclusion restriction. The 2008 fee abolition is deliberately *not* used as an instrument: it was itself a school-finance reform, replacing abolished fees with transfers, so it moves spending directly and fails exclusion by construction.

Both instruments are strong, and every estimate lies at or below the panel elasticity of 0.51. The demographic instrument gives 0.30 (s.e. 0.09), statistically indistinguishable from least squares on the same window (0.28) — no evidence that demographic enrollment variation is endogenous to budgets — while the policy instrument gives 0.04 (s.e. 0.15): enrollment changes induced by the access tightening, the margin closest to the model’s experiments, meet essentially *no* contemporaneous budget response. The gap between the recent-window estimates (near 0.3) and the full-panel 0.51 reflects the sample period, not the instruments, and sits in-

side the range the long differences already display (Table 2). Two considerations locate the policy-IV estimate within this range rather than at its center. It is an *impact* response — budgets observed within five years of the shock — while the model’s experiments run at generational horizons, and the long differences in Table 2 show fiscal adjustment building with horizon (0.32 at five years, 0.47 at ten): a near-zero impact response and a decadal elasticity near one half are the same sticky process seen at two horizons. And its literal reading, $\beta_g \approx 0$, is not outside the paper: it is the fixed-budget column reported throughout the policy analysis, in which the reform remains valuable (+0.46, or +15% of consumption) and the incumbent lineages, rather than gaining, pay for it through dilution. The reading for the model is thus one-sided: exogenous enrollment shocks meet budgets at least as sticky as the calibrated rule assumes, so the dilution mechanism is, if anything, understated at $\beta_g = 0.51$ — a choice that is conservative for the paper’s central mechanism and generous only to the claim that the top lineages gain, which Section 5 reports across the credible range.

Table A13: The fiscal elasticity: instrumented estimates of β_g

	panel OLS	demographic IV window (2006–16)		policy IV window (2010–19)	
	1990–2023	OLS	IV: predicted cohorts	OLS	IV: exposure $\times$ post-2014
log K–12 enrollment	0.512 (0.040)	0.284 (0.045)	0.301 (0.086)	0.296 (0.044)	0.039 (0.150)
first-stage F	—	—	220	—	54
observations	6,058	3,426	3,426	2,864	2,864

Notes: Two-stage least squares on specification (4) (log city education spending on log K–12 enrollment, controlling log total fiscal expenditure and college enrollment, city and year fixed effects; standard errors clustered by city). Demographic instrument: log school-age population predicted from the prior census’s city age pyramid shifted forward (2005 mini-census for 2006–11, 2010 census for 2011–16; city intercepts census-window specific). Policy instrument: 2005 migrant share of adults $\times$ post-2014. Using both instruments jointly on the demographic window gives 0.261 (0.085). The 2008 fee abolition is excluded as an instrument because it changed school finance directly.

E.9 The co-residence target: denominator reconciliation

The calibration’s co-residence target of 0.37 is a stock ratio built on a consistent *rural basis* from 2010-census counts: 35.8 million children of rural–urban migrants living with their parents at the destination, against 61 million rural left-behind children ([All-China Women’s Federation, 2013](#)). Both counts refer to children aged 0–17 of parents holding agricultural hukou; “migrant” means residing outside the registration county; a child is left behind if at least one parent is a migrant and the child resides in the registration county. The target is $35.8/(35.8 + 61) = 0.37$: among children of rural migrant parents, the share living at the destination. It is the complement of the widely cited left-behind share within the same population, uses the same reference date, and excludes urban-to-urban migrant children (whose model counterpart faces the same enrollment schedule but is a small population). The model moment is constructed identically:

brought children of rural-lineage migrant parents over all children of rural-lineage migrant parents. One definitional asymmetry is inherited: the published left-behind count includes children with one parent still present, while the model’s left-behind state has no co-resident parent, so the target understates co-residence in the model’s sense; the fitted settlement cost m^c is correspondingly higher, which dampens the reform’s reunification response — a conservative direction for the paper’s results. A both-parents-absent alternative cannot be constructed from the published stock counts, which do not split the left-behind total by the number of absent parents.

F The access schedule: sources and construction

The benchmark schedule $\{\rho_i\}$ of Table 3 is constructed from published public-school enrollment shares of migrant children, primarily [Chen et al. \(2020\)](#), whose figures refer to 2014: a national share of 79.5%, and elementary-level shares in the largest destinations of 41% (Guangzhou), 46% (Shenzhen), and near 70% (Beijing and Shanghai). Tier values are assigned as follows. The megacity tier takes the migrant-population-weighted average of the four first-tier cities’ shares, ≈ 0.57 . Tier 6 is set to 0.70 and tier 5 to 0.75, interpolating between the first-tier cities and the national average in line with the qualitative ranking of restrictiveness by destination size in [Chen et al. \(2020\)](#) and in enrollment-barrier rankings from ongoing work ([Guo et al., 2026](#)). Tiers 2–4 are set to 0.86–0.90, above the national average, consistent with near-parity in small destinations (the national share averages over the restrictive large cities). The rural tier is $\rho_1 = 1$: a child schooled in its registration county faces no enrollment barrier by construction. Finally, the schedule is renormalized (a factor of 1.019) so that the migrant-child-weighted mean equals 0.70, preserving the aggregate enrollment level implied by the national share and the model’s own distribution of migrant children across tiers. Two caveats are inherited by everything downstream. The tier values are enrollment *shares* used as enrollment *probabilities* — an equilibrium outcome standing in for a policy object, which conflates admission rules with application behavior and selection; and the 2014 vintage postdates the 2005 population distribution of the benchmark, so the calibrated economy is a synthetic benchmark rather than a dated equilibrium. The reform-based event studies of Appendix E.7 supply the response half of the program that would replace this schedule — family separation moves causally with access policy; what remains is a policy-based access *measure* (from eligibility rules, documentation requirements, and points thresholds) to serve as the dose. The single-tier bracketing of Section 7 and Appendix G spans the range such a measure would plausibly occupy, and Table A14 brackets the *full* schedule: it re-solves the model and the equal-access reform from an early-2000s schedule (the benchmark scaled to the pre-liberalization national share of about 0.5) and from a post-2014 schedule (mandated relaxation below, points systems in the top tiers at the dose of the 2014 episode). The paper’s headline is the conservative end of the range: both alternative vintages carry more megacity exclusion than the benchmark, so the reform’s gain is larger from

either, every lineage gains throughout, and the mobility effect grows with the removed risk. No conclusion of the paper depends on the schedule’s vintage within its plausible range.

Table A14: The equal-access reform under alternative access schedules

	early-2000s	benchmark (2014 shares)	post-2014
Schedule, tiers 2–7	.64/.63/.61/.54/.50/.41	.90/.88/.86/.75/.70/.57	.95/.95/.95/.85/.55/.42
Excluded share, brought migrants	0.42	0.24	0.27
Co-residence rate, rural migrants	0.32	0.375	0.35
Δ entrant welfare, net of tax (CEV)	+1.29 (+47%)	+0.84 (+28%)	+0.96 (+33%)
ΔW_7 , megacity-born, net of tax	+0.40	+0.26	+0.22
ΔW_1 , rural-born, net of tax	+2.01	+1.29	+1.48
IGE, benchmark $\rightarrow$ reform	0.64 $\rightarrow$ 0.45	0.555 $\rightarrow$ 0.433	0.60 $\rightarrow$ 0.46
Co-residence response (pp)	+10.2	+5.1	+6.6
Tax rate τ , long run (%)	9.0	8.6	8.8

Notes: Each column solves the stationary benchmark at the indicated schedule (all other parameters unchanged; the fiscal rule anchored at that benchmark’s own enrollment) and the $\rho_i = 1$ reform from it, under the estimated fiscal rule with the tax cost charged. The early-2000s column scales the benchmark’s urban schedule to the pre-liberalization national enrollment share of about 0.5; the post-2014 column relaxes tiers 2–5 and sets tiers 6–7 to the 2014-episode dose. Benchmark-fit rows shift with the schedule since no parameter is re-fit (co-residence falls as exclusion risk rises); both alternatives carry more exclusion in the largest tiers than the benchmark, which is why the reform’s gains exceed the headline from either end.

G Single-tier incidence: the full scan

This appendix completes the incidence exercise of Section 7 for every tier. Figure 4, panel (a), varies the small tier’s access alone. The nation and the rural-born lineage prefer it open: national welfare rises monotonically in ρ_2 , as it does in every tier’s access. But the tier’s *own* lineage’s welfare falls monotonically, all the way to the corner. Closing its schools to migrant children entirely is worth +12% of consumption ($\Delta W = +0.38$) to the tier-2-born, at a national cost of 4.4% (0.15); the lineage holds 0.6% of births, so the gain is large per head and negligible in the aggregate. Two features of the model deliver the corner. First, exclusion inflicts a bounded loss: an excluded child is still schooled, at family expense, so the harm per excluded child is the finite gap between the enrolled and excluded values. Because the harm is bounded, migrant families are never deterred entirely, tightening keeps paying the incumbents at every level of restriction, and nothing stops their preferred policy short of full closure. Second, in a small tier the fiscal arithmetic favors exclusion: the native child base is tiny — few children are born there — so even a modest migrant inflow looms large against it, and each excluded migrant child raises spending per remaining child materially, while the tier is too small for the lost migrants to dent its tax base. A lineage-welfare constituency for restriction therefore exists in the model — and Table A15 completes the scan for every tier, reporting each lineage’s stake in its own tier’s access beside its population weight. The own-lineage gain from full closure falls steeply with tier size: +12% of consumption ($\Delta W = +0.38$) for the tier-2-born, +2.7% (+0.09) for

tier 3, at most +1.8% (+0.06) in tiers 4–6, and ± 0.01 — nothing — in the megacities. Where such a gain exists it is held by small lineages (tiers 2–3 hold 4.9% of births between them), and everywhere the nation’s loss from closure exceeds the incumbents’ gain by growing multiples of the tier’s size.

Table A15: Own-lineage and national stakes in single-tier access

	Urban tier					
	2	3	4	5	6	7 (mega)
Own-lineage gain from full closure ($\rho_i \rightarrow 0$)	+0.38	+0.09	+0.01	+0.04	+0.06	+0.01
National cost of closure	−0.15	−0.08	−0.11	−0.22	−0.50	−0.56
Lineage birth share (%)	0.6	4.3	24.2	7.6	4.8	3.7
National gain from opening ($\rho_i \rightarrow 1$)	+0.02	+0.02	+0.02	+0.09	+0.24	+0.46
Rural-born gain from opening	+0.03	+0.03	+0.04	+0.15	+0.37	+0.70

Notes: Steady states under the estimated fiscal rule, one tier’s access moved to the indicated corner with every other tier at its benchmark value; entries are entrant-welfare changes relative to the benchmark at the experiment’s allocation (the tax adjustment of a single-tier experiment, at most 0.015, is common to all lineages and does not affect the comparisons). Tiers 2 and 7 from the full scans behind Figure 4; tiers 3–6 from corner solves. The own-lineage entry is each tier’s own-born lineage’s gain from full closure of its own schools to migrant children; rural-born gain is ΔW_1 from opening the indicated tier fully.

H Robustness

Robustness: the enrollment process. The benchmark draws enrollment i.i.d. across childhood periods; institutionally, admission may instead be decided once and persist. We therefore re-solve the model under the polar alternative: a family that stays in the same city with its child keeps the child’s enrollment status for as long as it stays — admission decided once per school career — with a fresh draw only on arrival in a (new) city. Three results. The *policy response is invariant to the enrollment process*: the model’s response to a 2014-style tightening (the frozen-price experiment of Section 4.3) is +4.1 percentage points of separation under the i.i.d. draw and +4.2 under full persistence at the same dose, and the two remain within a tenth of a point at double the dose — the family-arrangement elasticity that drives the paper’s experiments does not rest on the i.i.d. assumption. The *benchmark moves modestly*: at unchanged parameters, persistence lowers migration by five percentage points and co-residence by one, as the one-shot career lottery deters risk-exposed moves. And the *reform conclusions survive*: equal access under full persistence delivers a 25% consumption-equivalent welfare gain before financing ($\Delta W = +0.74$, against the benchmark’s +0.86 on the same basis; tax charges are nearly identical across the two specifications), every lineage again gains (megacity-born +5.8%; $\Delta W = +0.19$), reunification is the same size (+5.2 points), and the IGE falls by 0.11. Persistence slightly dampens the gains — under an absorbing draw, an excluded child stays excluded until the family moves, so removing the lottery is worth a little less to families that

would have re-rolled — but changes no conclusion.

Robustness: the warm glow. The contemporaneous parental valuation η could be suspected of manufacturing the family-arrangement response. It does not. Re-solving the entire model at $\eta = 0$, at half, and at double the benchmark value — re-fitting the settlement cost to the co-residence target in each case ($m^c = 1.14, 1.30, 1.85$ against the benchmark 1.46) — leaves every conclusion intact. The response to the 2014-style tightening episode is +3.6 percentage points *with the glow shut off entirely*, against +4.1 at the benchmark and +4.3 at double strength. The equal-access reform’s welfare gain and mobility effect are essentially unchanged across the grid: on the fixed-budget basis the gain is a 15–17% consumption equivalent throughout ($\Delta W +0.52, +0.50, +0.46, +0.50$ from $\eta = 0$ to double), and the IGE falls by 0.08–0.09 throughout. And the co-residence *level* is attainable at an ordinary settlement cost without any glow. What the family-arrangement margin requires is the discrete enrollment risk itself — exclusion is a first-order event for the family budget, with or without a paternalistic preference — while the glow’s quantitative role is to price the education-expenditure share, its calibration target, and to add roughly a tenth to the policy response. Together with the glow-excluded welfare criterion reported throughout, neither the behavior nor the welfare conclusions of the paper ride on the paternalistic term.

Robustness: where the gain comes from. A welfare gain of a quarter of consumption invites the suspicion that it is manufactured by the model’s amplifiers rather than by the access margin itself. Table A16 therefore shuts down each remaining amplifier one at a time — re-solving the benchmark at the changed parameter and rerunning the equal-access reform under the estimated fiscal rule, with the tax cost charged as everywhere else. Removing agglomeration ($\lambda = 0$) trims the gain from a 28% to a 25% consumption equivalent ($\Delta W +0.84$ to $+0.74$): scale economies contribute about an eighth. Removing the city human-capital spillover ($\delta = 0$) *raises* the gain to 38% ($\Delta W = +1.07$): the spillover partially insures excluded children, who absorb city human capital outside the school system, so without it school access is the whole of a city’s human-capital advantage and exclusion costs more — the gain rests on the schooling channel, not on the externality. Halving dynastic altruism ($\gamma = 0.25$) scales the gain with the weight on descendants, as it must for a reform whose returns accrue to children and grandchildren, but what remains is still a 16% consumption equivalent ($\Delta W = +0.45$) — with the megacity lineage flat rather than losing and the mobility gain intact (the IGE still falls by 0.10). No single mechanism carries the result: with agglomeration, the spillover, or half the altruism removed — and, from the preceding paragraph, with the warm glow removed — the reform remains large, lineage-Pareto, and mobility-improving. What cannot be removed is the enrollment risk itself.

Table A16: The equal-access reform with each amplifier shut down

	benchmark	no agglomeration $\lambda = 0$	no spillover $\delta = 0$	half altruism $\gamma = 0.25$
Δ entrant welfare, net of tax (CEV)	+0.84 (+28%)	+0.74 (+25%)	+1.07 (+38%)	+0.45 (+16%)
ΔW_7 , megacity-born	+0.26	+0.12	+0.29	+0.02
ΔW_1 , rural-born	+1.29	+1.17	+1.60	+0.70
IGE, benchmark $\rightarrow$ reform	0.56 $\rightarrow$ 0.43	0.55 $\rightarrow$ 0.44	0.59 $\rightarrow$ 0.44	0.67 $\rightarrow$ 0.57
Co-residence response (pp)	+5.1	+4.8	+6.0	+4.9
Tax rate τ , long run (%)	8.6	8.6	8.7	8.4

Notes: Each column re-solves the stationary benchmark at the indicated parameter (all others unchanged, no re-fitting) and runs the $\rho_i = 1$ reform from it under the estimated fiscal rule anchored at that benchmark's own enrollment; welfare changes are net of the balanced-budget tax cost, in dynastic utils with consumption-equivalent variation in parentheses (computed at each column's own dynastic discount factor). Benchmark moments drift where the removed mechanism was doing calibration work: without re-fitting the settlement cost, halved altruism lowers benchmark co-residence to 0.19 (target 0.37), so its column understates the family-arrangement stake. The remaining amplifiers are shut down elsewhere: the warm glow in the preceding paragraph, the financing rule in Table 10, private instruction in Appendix C, and the access schedule itself in Appendix F.

Robustness: the scale-setting assumptions. Four modeling choices set the scale of the central mechanism rather than amplify it, and Table A17 varies each — re-solving benchmark and reform at the changed assumption, all other parameters fixed (benchmark-fit moments shift where noted, since nothing is re-fit). *The excluded-state input:* the benchmark gives excluded children zero public input; granting them 5% or 15% of g_i (financed within the same budget) scales the results with the enrolled–excluded gap, as it must — and at 15%, three times any plausible funding of migrant-run schools, the reform keeps 78% of its welfare gain and the megacity lineage still gains. Conversely, applying the 0.9 out-of-district funding wedge to *enrolled* migrant children barely moves anything. *The conversion wage:* converting public instruction at a common national wage, instead of the parent's work-city wage, *raises* every response — the reform's gain to a 53% consumption equivalent ($\Delta W = +1.42$), reunification to +7.0 points — so the benchmark convention is the conservative one: deflating the home school by the destination wage makes leaving a child behind *less* costly for big-city migrants, dampening the reunification the reform unlocks. *Congestion:* at a net elasticity of -0.05 — gross agglomeration more than undone by urban costs — the reform keeps 83% of its gain and the mobility effect is intact; what changes is the megacity lineage, whose gain shrinks to about zero, since the in-migration that rebuilds its budgets now carries a congestion cost. *The budget base:* recomputing s_E on compulsory-plus-secondary spending alone (0.06) keeps 71% of the gain. Finally, *the wage-bill elasticity:* the rule imposes a unit elasticity of the education budget with respect to the local wage bill — only the enrollment elasticity is estimated. Setting the wage-bill exponent to the estimated fiscal-expenditure elasticity (0.55) leaves the benchmark untouched but changes the recovery: megacity spending per enrolled child bottoms at -13% and settles 8% *below* its pre-reform level rather than 9% above, because the damped budget response passes less of the reform's tax-base growth back to the schools. The aggregate gain

remains +19% of consumption ($\Delta W = +0.58$) and the rural lineage's +34% (+0.97), but the megacity-born lineage slips to -2.6% (-0.09): like the enrollment elasticity's lower range, the wage-bill elasticity is a margin that decides the sign of the most protected lineage's small stake — and nothing else.

Table A17: The equal-access reform under the scale-setting assumptions, varied one at a time

	bench	excluded 5% g	input 15% g	wedge 0.9 g	national wage	$\lambda =$ −0.05	$s_E =$ 0.06	wage-bill elast. 0.55
Δ entrant welfare (before financing)	+0.86	+0.79	+0.67	+0.79	+1.42	+0.71	+0.61	+0.58
ΔW_7 , megacity-born	+0.29	+0.27	+0.24	+0.24	+0.49	+0.05	+0.14	−0.09
ΔW_1 , rural-born	+1.31	+1.21	+1.02	+1.22	+2.06	+1.12	+0.92	+0.97
Co-residence response (pp)	+5.1	+4.5	+3.7	+4.9	+7.0	+4.6	+4.1	+4.5
IGE fall	0.12	0.11	0.09	0.12	0.14	0.11	0.08	0.10

Notes: Each column re-solves the benchmark and the $\rho_i = 1$ reform at the changed assumption under the estimated fiscal rule, holding every calibrated parameter fixed; changes are steady-state entrant-welfare differences at the allocation (the financing charge, common within each column, does not affect within-column comparisons; the baseline column's net-of-charge value is +0.84). Excluded-input columns give excluded migrant children the stated share of g_i , financed within the same budget (the enrolled-equivalent headcount adds $\text{exc} \times (1 - \rho_i)$ per migrant child); the wedge column funds enrolled migrant children at 0.9 g_i . The national-wage column converts the instructional component of public spending at the economy-wide average wage. The wage-bill column sets the budget's exponent on $w_i N_i$ to 0.55 around its benchmark value, so the calibrated equilibrium is unchanged by construction. Unrefit benchmark moments shift moderately in the national-wage and s_E columns (co-residence 0.391 and migration 0.397 against targets 0.370 and 0.426).

I Computational appendix

Algorithm. The equilibrium is solved as nested fixed points. Innermost, the value functions. Because W appears inside v_c (through the graduation node) yet is itself defined from the entry values of v_c , the life-cycle solve is a *fixed point in W* : guess W , solve the life cycle backward using it at graduation, update W from the entry log-sum-exp, and iterate to convergence. The dynastic discount γ (net of the life-cycle discounting embedded in v_c) keeps this map a contraction at the calibrated altruism strength. On the distribution side, since birth city is conserved, the entrant marginal $\{\pi_i^{\text{birth}}\}$ is held fixed and only the within-lineage human-capital distribution is solved as a fixed point. Outermost, a multiplicative tatonnement on the city-wage vector $\{w_i\}$: given a wage guess, the value functions and stationary distribution are computed, city efficiency labor N_i is aggregated, and each w_i is updated toward $\bar{A}_i N_i^\lambda$ — with the city-average human capital $\bar{H}_i$ and the fiscal objects updated alongside — until joint convergence. The transition of Section 5 is solved in the same way in time-indexed form: given guessed paths $\{w_i(t), g_i(t), \bar{H}_i(t)\}$, a single backward sweep of the value functions and a single forward sweep of the distribution imply updated paths, iterated to convergence. The dynastic altruism resolves within the backward sweep — a child graduating at t becomes a labor-market entrant valued at the date- t entry value — so no separate generational fixed point is needed along the

path.

State space. Human capital is on an $n_h = 9$ -point log-spaced grid on $[0.1, 25]$ (log spacing follows the log-normal shape of the stationary human-capital distribution; Section 4.3 documents that linear or too-narrow grids clamp the right tail and artificially compress the dispersion and intergenerational moments). Financial wealth is on an $n_y = 4$ -point grid on $[0, 30]$ with curvature 2 (wealth plays a limited role: the borrowing constraint binds for most families with children, and savings are decumulated before the graduation handoff). The education shock is integrated with 3-point Gauss–Hermite quadrature; child human capital between grid points is handled by linear interpolation in logs. The with-child problem’s two continuous choices (child expenditure and saving) are solved by nested golden-section search (16 iterations each) with corner clamps; the parental money/time split within child expenditure has a closed-form cost-minimizing interior solution.

Grid refinement. Because the reported moments are distributional, Table A18 re-solves the stationary benchmark and the equal-access reform on two refined state spaces — $(n_h, n_y, n_q) = (15, 8, 5)$ and $(25, 8, 5)$ against the baseline $(9, 4, 3)$ — at the calibrated parameters. Every response the paper leans on is stable: the co-residence response, the fall in the IGE, the reform’s welfare gain, and the lineage changes move by less than a tenth of their values, with the gains slightly *larger* on the refined grids; benchmark dispersion edges toward its target (1.29 against 1.33 at the finest grid). The Section 7 megacity null survives refinement: on the $(15, 8, 5)$ grid the megacity-born lineage’s stake in its own tier’s access ranges from -0.002 (full closure) to $+0.031$ (full openness) — within $\pm 1\%$ of consumption, against the tier-2-born lineage’s closure gain of $+0.38$ — so the incidence conclusion is not a discretization artifact.

Table A18: Grid refinement: headline moments and responses across state spaces

(n_h, n_y, n_q)	(9, 4, 3) baseline	(15, 8, 5)	(25, 8, 5)
SD(ln income), benchmark	1.330	1.302	1.293
IGE, benchmark $\rightarrow$ reform	0.555 $\rightarrow$ 0.433	0.588 $\rightarrow$ 0.469	0.597 $\rightarrow$ 0.476
Excluded share, benchmark	0.242	0.242	0.243
Co-residence response (pp)	+5.1	+5.2	+5.2
Migration response (pp)	+4.6	+4.8	+4.7
Δ entrant welfare (before financing)	+0.86	+0.93	+0.92
ΔW_7 , megacity-born	+0.29	+0.31	+0.29
ΔW_1 , rural-born	+1.31	+1.36	+1.35

Notes: Each column solves the benchmark and the $\rho_i = 1$ reform at the indicated human-capital, wealth, and quadrature resolutions, all parameters at their calibrated values; the fiscal rule is re-anchored at each grid’s own benchmark enrollment, so the anchored-rule structure is preserved exactly. Welfare levels are grid-dependent; the reported objects are within-grid differences.

Numerical uniqueness. The education technology could in principle support multiple development equilibria (Section 4.3). As a check, the benchmark, the equal-access reform, universal conversion, and the two extreme single-tier closures (tier 2 and tier 7) were each solved from two initializations of $\bar{H}_i$ spanning the grid — the optimistic level used throughout and the grid

floor ($\bar{H}_i = 0.02\bar{h}$): every pair converges to the same equilibrium, with welfare differences between 2×10^{-5} and 1.1×10^{-4} , inside the equilibrium tolerance of 10^{-4} . The equilibrium is therefore numerically unique over the reported policy range — a numerical statement over these checks, not an analytical theorem — and no reported comparison depends on initialization.

Fixed points and tolerances. The value-function block solves the W -fixed point described in Section 4 to a sup-norm tolerance of 10^{-8} (the map contracts at rate $\approx \gamma\beta^{J_c} = 0.17$ per iteration). The outer spatial equilibrium iterates a damped tatonnement (damping 0.9) on the wage vector jointly with $\{g_i, \bar{H}_i\}$ to a relative sup-norm tolerance of 10^{-4} ; the benchmark converges in 24 outer iterations with terminal error 8.6×10^{-5} , and all reported moments are stable well before the tolerance binds. The transition solves backward–forward sweeps over a horizon of $T = 24$ periods (168 years) on guessed paths $\{w_i(t), g_i(t), \bar{H}_i(t)\}$, damped identically; the equal-access path converges in 28 outer iterations, and the terminal condition is the value function of the independently solved reform steady state, so households internalize the post-horizon future rather than a truncated one. At the last plotted period the fast-moving series have essentially reached that steady state (co-residence 0.4250 against 0.4256), while the slow-moving ones are still approaching it: megacity spending per enrolled child stands at +5.1% against +9.1% in the terminal steady state, and migrant children’s mean graduation human capital at 1.701 against 1.714 — tail convergence on the generational clock of the human-capital stock, which the transition figures mark with right-edge steady-state ticks. Doubling the horizon leaves every reported transition statistic — impact, trough, and the welfare of the cohorts alive at the reform — unchanged at the reported precision.

The balanced budget in computation. Each experiment’s allocation is solved with households taxed at the budget share s_E ; the tax-rate path $\tau(t) = \sum_i g_i(t)K_i(t)/\sum_i w_i(t)N_i(t)$ is computed from the solved aggregates, and its dynastic consumption cost (the D -weighted discounted sum of $\log[(1 - \tau(t))/(1 - s_E)]$) is charged to reported welfare. The formal basis is the following invariance property. *Claim: scale every net-of-tax wage rate by a common factor $\varkappa = (1 - \tau)/(1 - s_E)$, holding $\{g_i\}$ and the child inputs fixed. Then every household’s optimal consumption–saving plan scales by $\varkappa$ (entrants hold no assets and the budget is linear, so the problem is homogeneous), every conditional value shifts by the same constant $D \log \varkappa$, and every discrete choice — location, bring, settlement, all Gumbel comparisons — is exactly unchanged, because log utility turns the common scaling into a level shift that cancels from every value difference. The welfare change is exactly $D \log \varkappa$: the charge.* The single margin outside the claim is the money input to children, whose re-optimization gain is second order by the envelope theorem, bounded by $\frac{1}{2}[(\tau - s_E)/(1 - s_E)]^2$ of consumption — below 0.02% at the largest tax gap in the paper (1.8 percentage points, the $\beta_g = 1$ column) — but whose level response feeds descendants’ human capital, a general-equilibrium term the charge does not price. We therefore verify the treatment with one complete stationary solution with the endogenous rate inside the household problem: the equal-access steady state re-solved with households pay-

ing $\tau^* = 8.65\%$, the rate its own budget implies. The solve confirms the claim exactly where it applies: the implied rate reproduces the imposed one to half a basis point (a fixed point), and choices are numerically invariant (co-residence moves 0.05 percentage points, migration 0.09, the IGE by less than 10^{-3}). It also prices the omitted feedback: lower net wages scale down money inputs, and aggregate human capital settles about 1% lower, so the complete solve delivers a reform gain of +0.81 against the +0.84 the accounting treatment reports. The charged numbers thus overstate net gains by roughly 3% of themselves, uniformly across lineages; no ranking, sign, or conclusion moves. Transitions retain the accounting treatment: solving the tax path inside the household problem along a transition destabilizes the fixed-point iteration itself: with the tax fed back into budgets mid-iteration, a small perturbation can carry the solver to a different fixed point of that iteration — unlike the anchor-rate solves reported throughout, which are numerically unique across initializations (see the uniqueness check below).