Spatial Inequality of Opportunity in West Africa. ONLINE APPENDIX

Luke Milsom

2nd April, 2023

Contents

1 Robustness

1.1 Linear model

Building on the linearity of place effects, I estimate a parametric version of my main equation allowing a different slope and intercept for children moving between the ages of 1 and 12 and those moving after age 12:

$$\begin{aligned} y_i = \rho_{os} + \alpha_m \\ &+ (\alpha_1 + \beta_1 \times \text{age move}_i) \cdot \mathbb{1}\left[\text{age move}_i \in [1,12]\right] \cdot \Delta\bar{y}^p_{ods} \\ &+ (\alpha_2 + \beta_2 \times \text{age move}_i) \cdot \mathbb{1}\left[\text{age move}_i \in [13,18]\right] \cdot \Delta\bar{y}^p_{ods} + u_i \end{aligned} \tag{1}$$

Table 1 shows the results from estimating equation 1. The rate of convergence, β_1, is as found in the non-parametric specification. Additionally, the post-12 slope is a precisely estimated 0. The first column of table 1 estimates the baseline specification, and columns 2, 3, and 4 estimate similar equations with varied fixed effects specifications. Column two replaces the cohort (year born) by birth location fixed effects with individual cohort fixed effects, allows cohort-specific slopes on birth location quality $\bar{y}^p_{os}$, and allows age at move fixed effects to vary by census. Column three removes the more flexible at age of move fixed effects of column 2. Finally, column four removes the year by $\bar{y}^p_{os}$ fixed effects and so no longer controls for origin location in any way. In all specifications, I find little change in results from the main specification given in column one.

Table 1 Parametric estimates of place effects

	Year born - birth location and age at move FE	Year born, year born by $\bar{y}_{os}$ and age move all by sample FE	Year born, year born by $\bar{y}_{os}$ and age move FE	Year born and age moved FE
$\hat{\beta}_1$	-0.0270***	-0.0272***	-0.0284***	-0.0319***
	(0.00530)	(0.00500)	(0.00554)	(0.00729)
$\hat{\beta}_2$	0.00302	0.00893	0.00553	-0.00742
	(0.00914)	(0.00631)	(0.0120)	(0.0234)
N	79778	79956	79970	79970
R^2	0.348	0.311	0.303	0.155

Note: This table shows the results from estimating equations of the form $y_i = FE + (\alpha_1 + \beta_1 \times \text{age move}_i) \cdot \mathbb{1}[\text{age move}_i \in [1,12]] \cdot \Delta\bar{y}^p_{ods} + (\alpha_2 + \beta_2 \times \text{age move}_i) \cdot \mathbb{1}[\text{age move}_i \in [13,18]] \cdot \Delta\bar{y}^p_{ods} + u_i$, where FE are the fixed effects specification which varies by column. In column one I include year born by birth location and age at move fixed effects (baseline specification); in column two I include year born, (slopes of) year born by origin location quality and age at move by sample fixed effects; in column three I include year born, (slopes of) year born by origin quality and age at move fixed effects; in column four I include year born and age at move fixed effects. Standard errors clustered at origin and destination.

1.2 Robustness of the LPM

In the main text I implicitly assume a linear probability model when performing the non-parametric or linearly parameterized regressions. In this section, I consider whether results substantively change if instead an alternative functional form is considered. In table 2 I replicate the results from table 1 in the main body in columns 1 and 2, where *few fixed effects* refers to year born, year born by locality quality, and age at move fixed effects. Columns 1 and 3 of table 2 use the full fixed effect specification of year born by birth location and age at move fixed effects. The other columns use the few FE specification to mitigate incidental parameter bias known to be pervasive in probit and logit settings. The coefficients all have the same sign and tell the same story of an exposure effect up to age 12 and no impact thereafter. Of course, one cannot directly compare coefficients across each type of model. PPML approaches in columns (3) and (4) estimate a semi-elasticity and so can be converted into percentage point effects (as with the linear probability model) by multiplying by the mean of the left-hand side variable, which in the estimating sample is 0.64. This results in very similar estimates to those in columns (1) and (2). As the logistic and normal distributions are similarly shaped up to a scaling factor, it is known that logit

Table 2 Robustness of the main results to the linear probability assumption

	(1) LPM	(2) LPM	(3) PPML	(4) PPML	(5) Probit	(6) Logit
1 to 12 slope	-0.0270***	-0.0283***	-0.0452***	-0.0478***	-0.0871***	-0.145***
	(0.00530)	(0.00545)	(0.0113)	(0.0121)	(0.0113)	(0.0189)
13 to 18 slope	0.00302	0.00469	0.00210	0.00951	0.0220	0.0295
	(0.00914)	(0.0119)	(0.0185)	(0.0248)	(0.0246)	(0.0418)
1 to 12 constant	0.486***	0.514***	0.856***	0.950***	1.557***	2.574***
	(0.0496)	(0.0611)	(0.122)	(0.133)	(0.178)	(0.292)
13 to 18 constant	0.0932	0.0805	0.208	0.174	0.135	0.339
	(0.122)	(0.166)	(0.238)	(0.346)	(0.367)	(0.624)
Year born by birth location and age moved FE	X		X			
Year born, year born by birth location quality and age moved FE		X		X	X	X
R^2	0.348	0.304				
N	79778	79778	78950	79776	79776	79776

Note: This table shows the results of re-estimating the linearly parametized specification relaxing the linear probability assumption. Column one shows the baselines results using a linear probability model and the full fixed effects. Column two also uses a linear probability model with *easy* fixed effects. Columns three and four use psuedo poisson maximum likelihood estimation with the full and easy fixed effects respectively. Column five estimates a probit model with easy fixed effects and similarly column six estimates a logit model.

coefficients will roughly be 1.82 times the probit coefficients - which is indeed what we find. Logit coefficients can in turn be interpreted as the percent impact on the odds ratio which is on average 1.78.

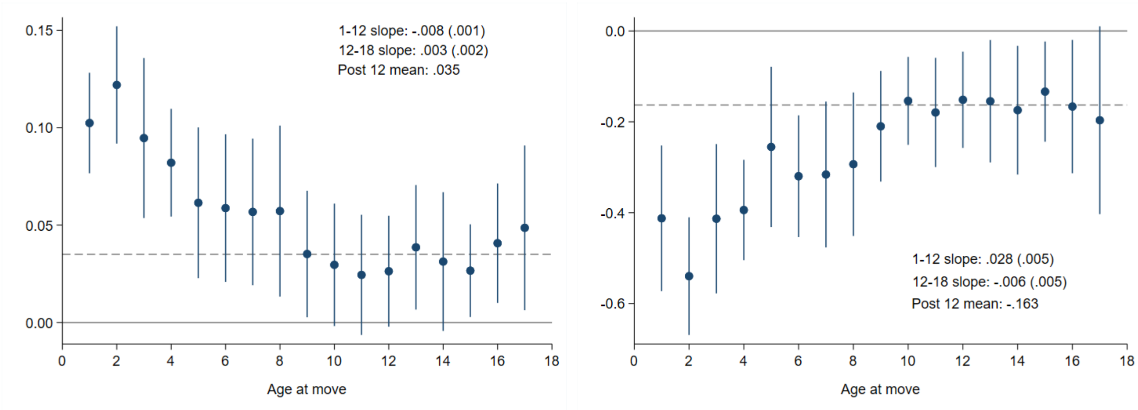
1.3 Considering different measures of location quality

In the above results I've used primary completion as the outcome variable and average primary completion among permanent residents as the measure of location quality. It's also interesting to consider whether different measures of location quality generate similar results. In figures 1a and 1b I show analogous results to the main result presented in figure **??** where instead of average primary completion among permanent residents I consider average housing quality and employment in agriculture among permanent residents respectively as my measure of location quality (the outcome variable remains primary completion as before).

This analysis serves two purposes. First, it acts as a robustness/ sense check - indeed given that I find that location quality when measured in terms of primary completion matters it would be very surprising if place measured differently didn't

Figure 1 Causal place effects in other variables

(a) Housing quality (b) Agricultural employment



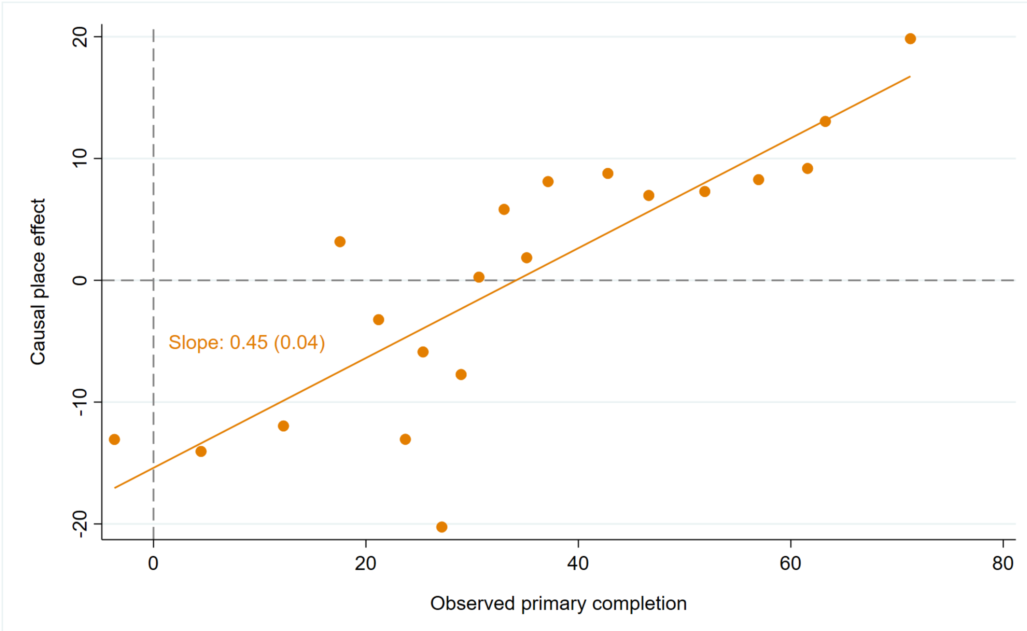
Note: This figure replicates the main results in the paper using different measures of location quality. Instead of measuring location quality as the primary completion rate among permanent residents, in the left panel I use housing quality and in the right agricultural employment (both for permanent residents). The dependent variable remains primary completion in all cases.

matter. In this way this exercise serves as a robustness check on my main results. It's also reassuring to find similar patters, with convergence up until age 10 to 12 or so and a constant selection effect thereafter.

2 Relationship between observed primary completion rates and estimated causal place effects

A simple sense-check on the estimated place effects is to study their relationship with observed primary completion rates. Figure 2 shows a (population-weighted) binscatter with the corresponding line of best fit both of which are residualized on census-level fixed effects. There is a strong positive and tight relationship between observed primary completion rates and causal place effects with a slope coefficient of 0.45. Figure 11 maps the estimated place effects.

Figure 2 Relationship between observed primary completion rates and estimated relative place effects



Note: This figure shows the binscatter relationship between the observed primary completion rates and estimated causal place effects (which have been multiplied by 13 to put into *moved at birth* terms). Census fixed effects are included, as are population weights. Standard errors are clustered at the locality level.

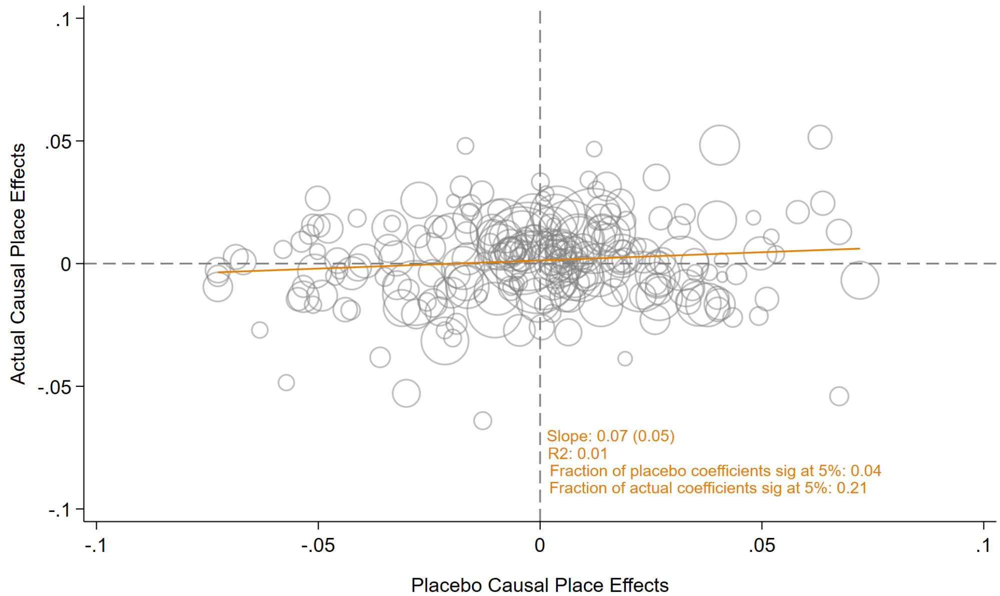
3 Locality-year specific identifying assumption

When estimating locality-year specific fixed effects I require a stronger identifying assumption: that selection effects are constant over age at move for each location l in each period t, rather than just on average.

Assumption 1 *Place specific identifying assumption:* $\mathbb{E}[e_{il} \cdot \varepsilon_i | \alpha_{odt}] = 0 \;\; \forall l \in \mathbb{L}$.

The results supporting the validity of the on-average identifying assumption summarized in the main text also provide evidence for this location-specific assumption. In addition, I perform a placebo test estimating causal place effects on a sample of those who move between 14 and 18, *after* primary education completion is determined. In this case, place effects μ_l should just be noise and unrelated to the actual estimated place effects. Figure 3 summarizes the relationship between actual and placebo place effects, statistics are weighted by locality population. As is clear from this figure they exhibit little relationship with a regression slope coefficient indistinguishable from zero and a very small R-squared. In addition, less than five percent of the individual placebo coefficients are significant at the 5% level.

Figure 3 Placebo vs actual causal place effects



Note: This figure shows the scatter plot relationship with a corresponding linear line of best fit between the actual and placebo-estimated place effects. Place effects in both instances have been residualised on year fixed effects. Circle size indicated population and the linear regression line has also been weighted by population.

4 Standard errors

In all individual-level regressions, standard errors are clustered at the origin locality and destination locality level (as in Chetty and Hendren [2018a,b]) allowing correlation within both units. In this section, I additionally consider a simple permutation-based test of the main linearized regression given in equation 1. I run 1,000 simulations where I randomly re-sort journeys, defined as birth-destination location pairs, across movers and re-estimate equation 1 on each new data set. That is, I do not permute between movers and stayers nor across the year of move, but instead across journeys within movers. This allows me to retain any bias in the data that may be stemming from the non-permuted dimensions i.e. mover vs stayer and year of the move. The resulting slope coefficients from 1,000 replications are plotted in figure 4 and show clear evidence of significant effects.

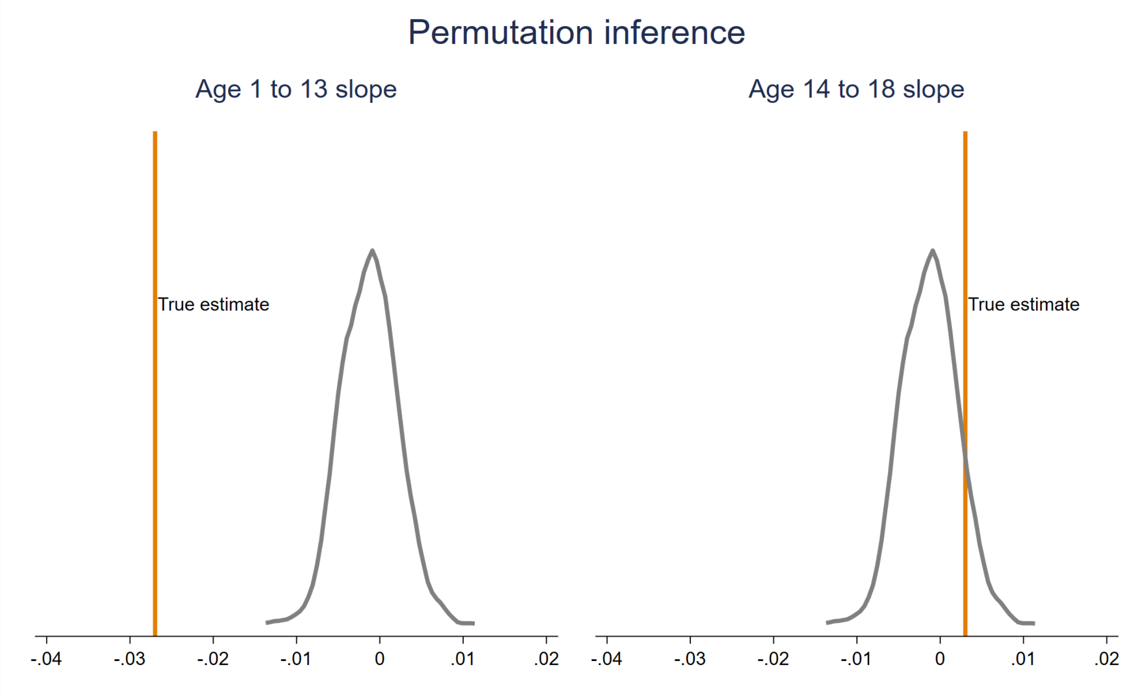


Figure 4 Permutation inference

Note: This figure shows the resulting distribution of estimated coefficients from a permutation test. I permute journeys (birth-destination pairs) across movers maintaining the observed year of move in 1000 replications estimating the linearised equation at each iteration.

5 Comparability of movers and stayers

Those who move, and those who stay, are not the same. Those who migrate are often more wealthy (as migration is costly) and may self-select, especially on the urban-rural dimension [Lagakos, 2020, Hamory et al., 2021] to areas where they can be most productive. Throughout this analysis, I use variation in movers to estimate the causal effect of place on primary education completion. To generalize results to the whole population, I implicitly assume areas that exhibit high place effects for movers would also do so for stayers. This does not mean that I assume movers and stayers must look the same, but more specifically that the impact of where a child grows up on their probability of completing primary education must be similar. It is not necessary to make this assumption, as the results would remain of interest if they only pertained to movers who comprise close to 20% of 14 to 18 year-olds. However, in this section, I consider whether the results apply more broadly to both movers and stayers.

The assumption that place effects are the same for movers and stayers can be decomposed into two components. First, that movers and stayers do not systematically differ in a manner that interacts with place effects. For example, it could be that movers are more entrepreneurial and therefore benefit relatively more from a high-quality location compared to stayers, and this is reflected in their propensity to educate their children. Secondly, it could be that movers and stayers are similar but due to the nature of having moved to a new location, movers interact differently with their surroundings. For example, a mover in a new location does not have connections or localized knowledge that a permanent resident may have. Therefore, although stayers and movers may be geographically co-located, they could interact with very different aspects of society.

To consider the second possible difference, I return to the main specification given in equation **??** and its linear parameterized version in equation 1. Due to the inclusion of origin by cohort fixed effects, this specification only uses variation in destination quality enjoyed after the move has occurred. Instead, I replace these fixed effects with destination by cohort fixed effects. This new specification uses variation in *origin* quality which movers enjoy *before* they moved. If the place effects I estimate when using variation in (to-be) movers' origin location quality are similar to those I obtain when using variation in their destination location quality, this is evidence that movers do not interact systematically differently with their new surroundings. Comparing convergence rates, I find that they are almost identical both statistically and economically. This implies that place exerts the same impact on those who are in their birth location and have yet to move as it does on those who have already moved and are residing in their destination location.

However, this result doesn't preclude the first concern, that movers and stayers are fundamentally different and that this interacts with place. To investigate this, I turn again to increasingly likely to be exogenous move events. In these cases, there is increasingly little selection into moving. Therefore, if I find similar causal place effects among the sample of more likely to be exogenous movers as I do in my overall sample, this will be evidence against the concern that movers are fundamentally different from stayers in a way that interacts with place effects. This is exactly what I find when performing this exercise in the online appendix.

I've highlighted two possible issues that can be ameliorated by considering increasingly likely to be exogenous moves: varying selection effects and the potential differences between movers and would-be stayers. It could be that these two effects are canceling each other out and that considering increasingly likely to be exogenous moves sheds no light on either issue. For example, suppose that selection effects are increasing in the child's age at move and that would be stayers (those who were it not for an exogenous move event, would not have moved) experience much weaker place effects than movers. These two fundamental issues would be hidden by just looking at increasingly likely to be exogenous move events. However, it is unlikely that selection effects vary with the child's age at move in this manner. It is more intuitive to suppose that households with better unobservables move earlier to better places. For the observed result to hold it would have to be that stayers exhibit much *stronger* place effects than movers which seems unlikely, but if so implies that my analysis is a lower bound for place effects. Given the additional validation checks performed above, I find it unlikely that either force is at play here but cannot rule it out entirely.

6 Descriptive statistics

6.1 Temporal variation in primary completion

Figure 5 shows the distribution of primary completion rates across localities for six countries in Sub-Saharan Africa, Benin Mali and Cameroon which are the main focus of this study as well as Tanzania, Zambia and Kenya for comparison[1]. Evident in all countries is a considerable rightward shift in primary completion rates from the early to the late period, although these changes are not directly comparable as the size of the time difference varies between countries. What is also striking is the heterogeneity between Benin, Cameroon, and Mali highlighting the stark differences between these countries as well as variation that can be exploited when I consider heterogeneity. Cameroon has a far more diffused distribution as well as both relatively high means and a relatively large mean shift. Mali on the other hand is more bunched at the lower levels of completion and displays a more modest shift. Benin shows considerable gains seemingly across the whole distribution.

Although not crucial for my analysis, it's helpful for the interpretation of results that indeed most individuals who will at some point receive primary schooling do so by aged 14. This is evidence from figure 6 which shows the age fixed effects from

[1]This sample was chosen purely for practical data availability reasons, Tanzania Kenya and Zambia are the only additional countries in Sub Saharan Africa which fill the data requirements, although sadly they don't have sufficiently rich migration information to be part of my main sample.

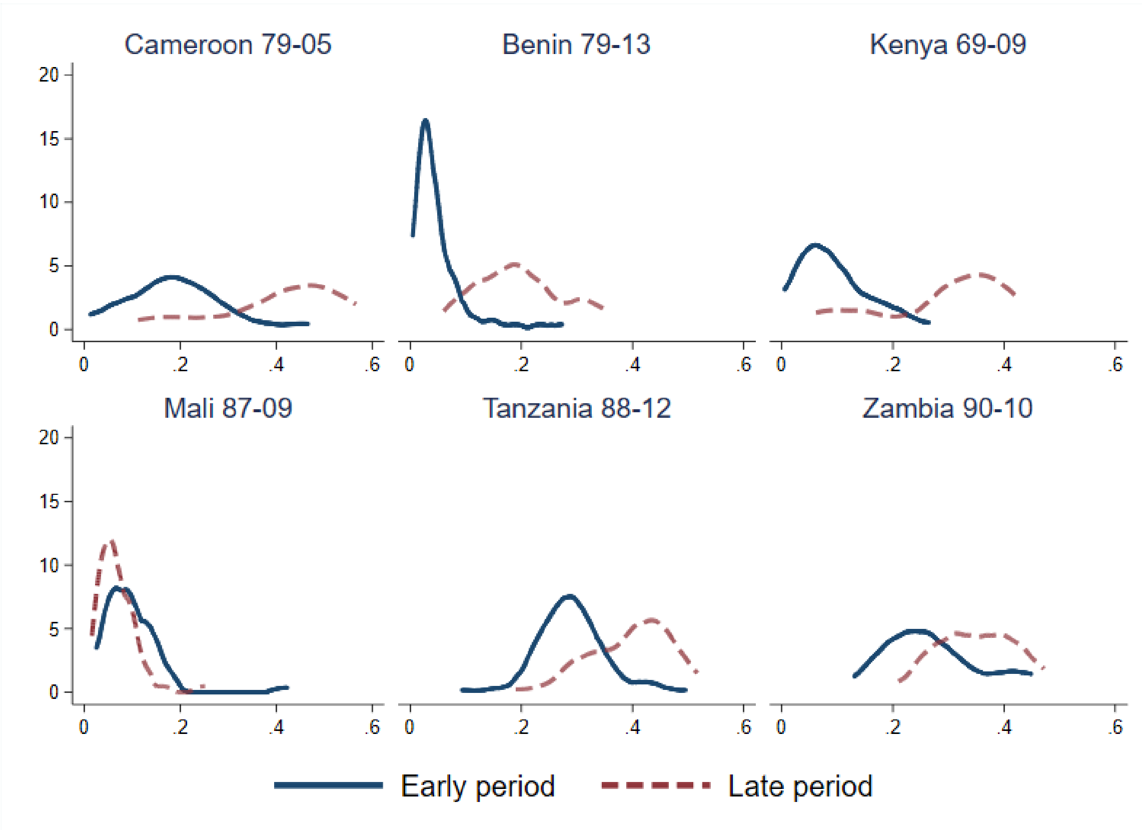


Figure 5 Change in the primary education completion distribution

Notes: This figure shows kdensity plots over locality level primary completion rates by country and year on a sample of over 12 year olds. In each case the solid blue line refers to the distribution in the earliest available census year in each country and the dashed red line the corresponding distribution in the most recently available census.

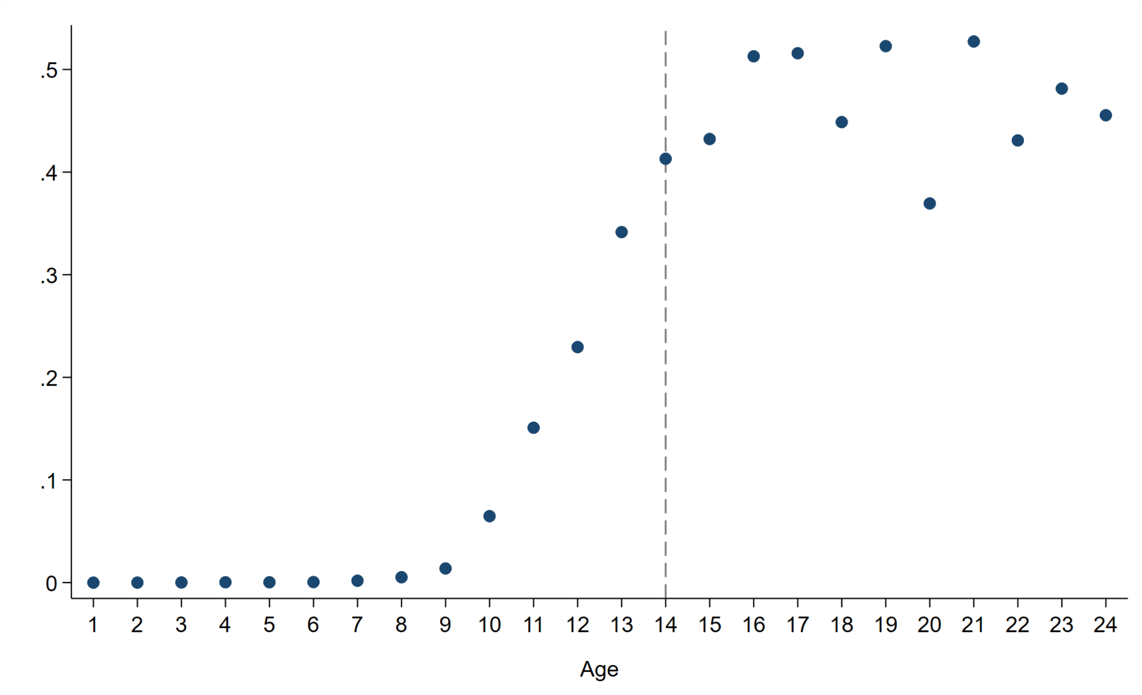


Figure 6 Proportion of the population who have completed primary school by age

Notes: This figure shows the proportion of the sample who report having completed primary education at the time of each census against their age. This figure uses data from the full sample Benin (1992, 2002, 2013) Cameroon (1976, 1987, 2005) and Mali (1998, 2009).

a regression of primary completion against age and year fixed effects. This figure clearly shows, as anticipated, that most primary education is indeed completed by 14. However, it's worth noting that many children who haven't completed primary education by 12, which is the year officially primary schooling ends in the countries I study, go on to do so in the next few years.

6.2 Sub-national migration patterns

The rich migration information collection in the censuses I use in this paper allows a detailed analysis of migration patterns at the sub-national level over a long time span. These patterns are important for my analysis because to estimate place effects I necessarily use variation in *movers* outcomes alone. If, for example, the majority of migration was towards the primate city I would have little external validity in consider counterfactual, or actual, changes in market access in the hinterland. Additionally, it would be challenging or impossible to estimate heterogeneity in effects across location characteristics. Figure 7 shows the proportion of migration directed

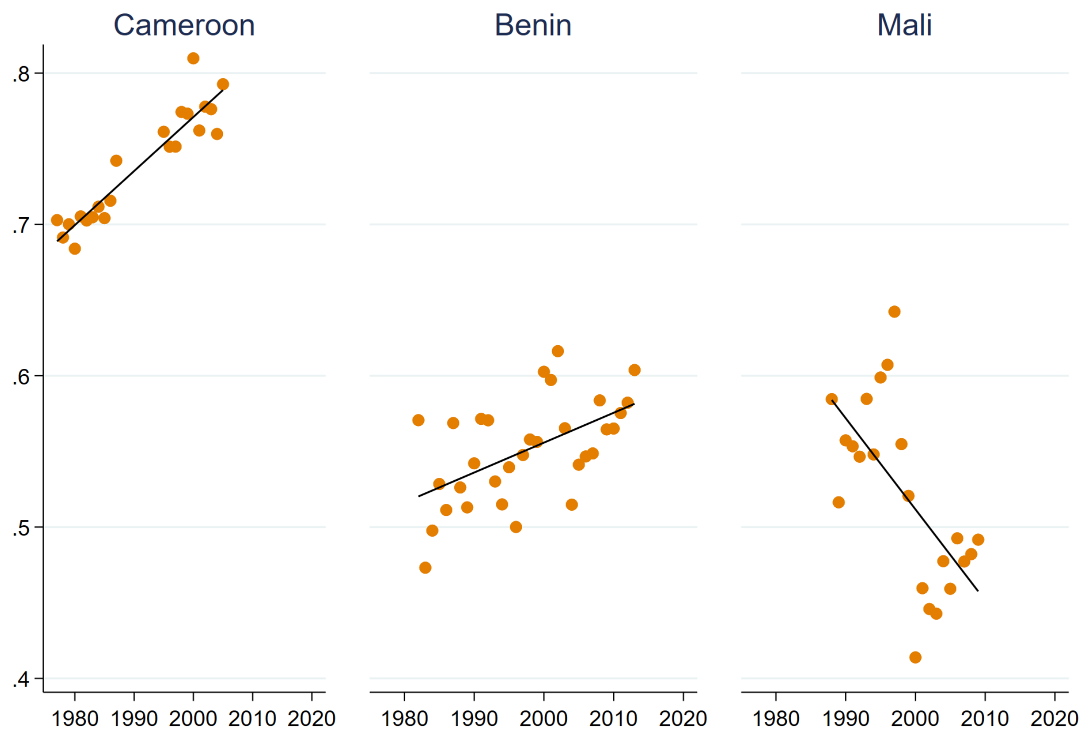


Figure 7 The proportion of migration to urban areas

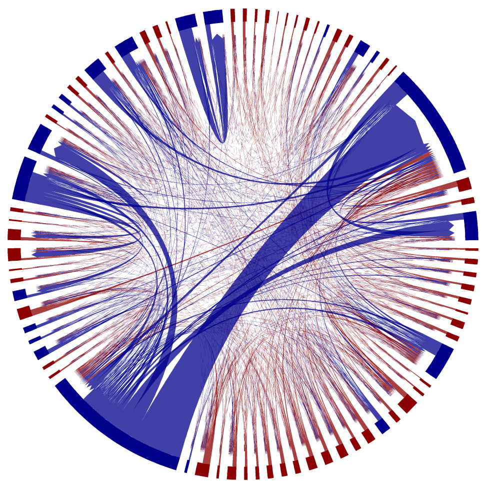
Notes: This figure shows the proportion of movers in a given year that move into an urban location where urban locations are defined as areas where over 50% of the residents report living in an urban locality. On each scatter plot a linear regression line has been added in red.

towards localities with over 50% of residents self-reporting as living in an urban area. There are considerable differences across countries and time, allying fears that we are only capturing one type of mover variation.

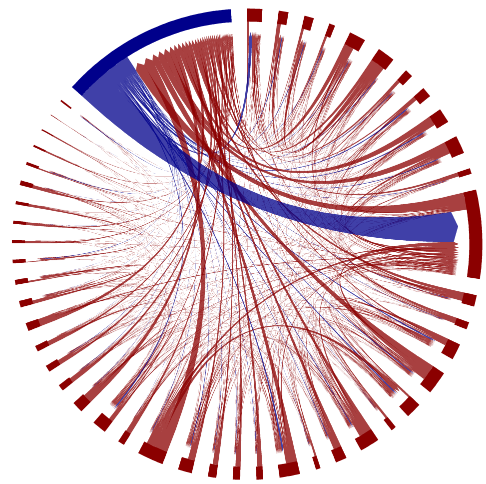
Additionally, figures 8a, 8b and 8c show flows in and out of each locality within each country over a 10 year period. Blue regions are classified as urban and red as rural. The size of the arc associated with each locality is proportional to the total movement in and out of the locality. One can see from these figures that sub-regional migration in Benin, Cameroon, and Mali cannot be easily characterized. There is primate to hinterland migration, hinterland to primate migration, urban to rural migration, rural to urban migration, urban to urban migration, and rural to rural migration. The objective of these figures is not to trace out-migration patterns specifically but rather to stress the overall heterogeneity in migratory decisions.

In Benin, migration is dominated by the two main cities, Porto-Novo (the larger

(a) Benin (2003-2013)



(b) Mali (1999-2009)



(c) Cameroon (1995-2005)

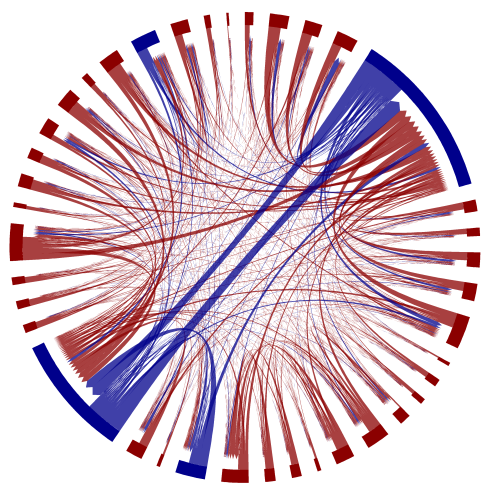


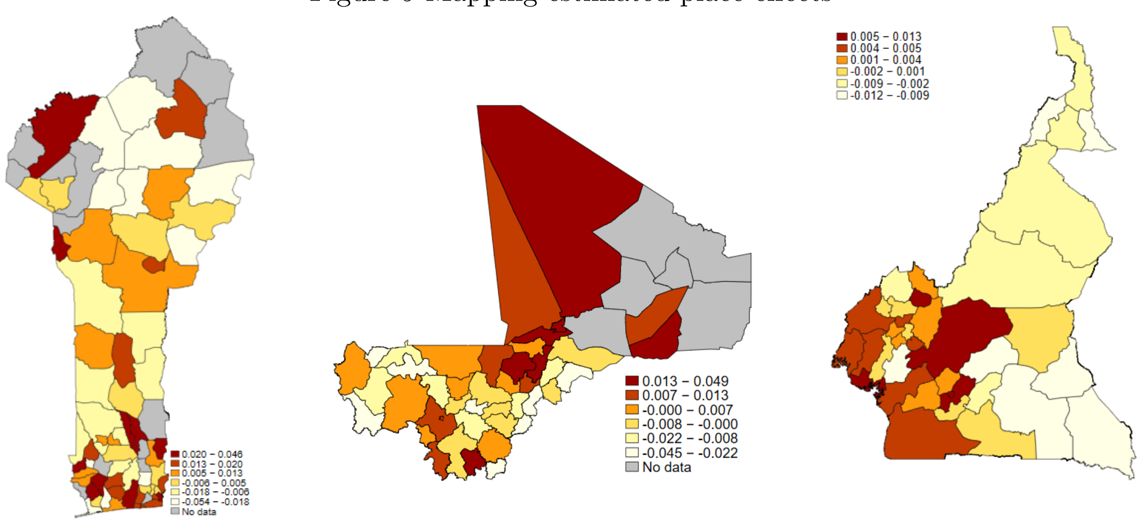
Figure 8 Migration chord diagrams

Notes: These figures are Chord diagrams showing population movements from and to each locality in my data within each country aggregating over the last observed 10 years. These figures will underestimate total migration as for individuals that move multiple times in the 10 year period only the most recent move is captured in my data. Each section of the circumference corresponds to a locality with size equal to the total in plus out migration observed. Arcs and flows colored in blue correspond to localities and movements out of said localities that are characterized as urban, that is greater than 50% of the individuals living in these localities report living in an urban area. Similarly red arcs and flows correspond to rural localities and movements.

band) and Cotonou. These cities are geographically very close to each other and a clear pattern of moving away from the capital to the more dynamic *up and coming* Cotonou. We also see a more general exodus from the capital and influx into Cotonou from other localities within Benin. Mali is the most primate-heavy country of the three, the capital Bamako (largest arc) has a population of almost 2 million compared to the next biggest city Sikasso (second biggest arc) where only 250,000 or so people live. Figure 8b shows that unlike Benin the capital of Mali shows far greater inflow than outflow attracting Malians from across the country. Cameroon is a tale of two cities, Douala (top right) and Yaoundé (bottom left) both home to about 2.5 million people. There is far greater inflow than outflow to these two areas, but that's not to say that there isn't significant movement not involving these locations.

7 Maps

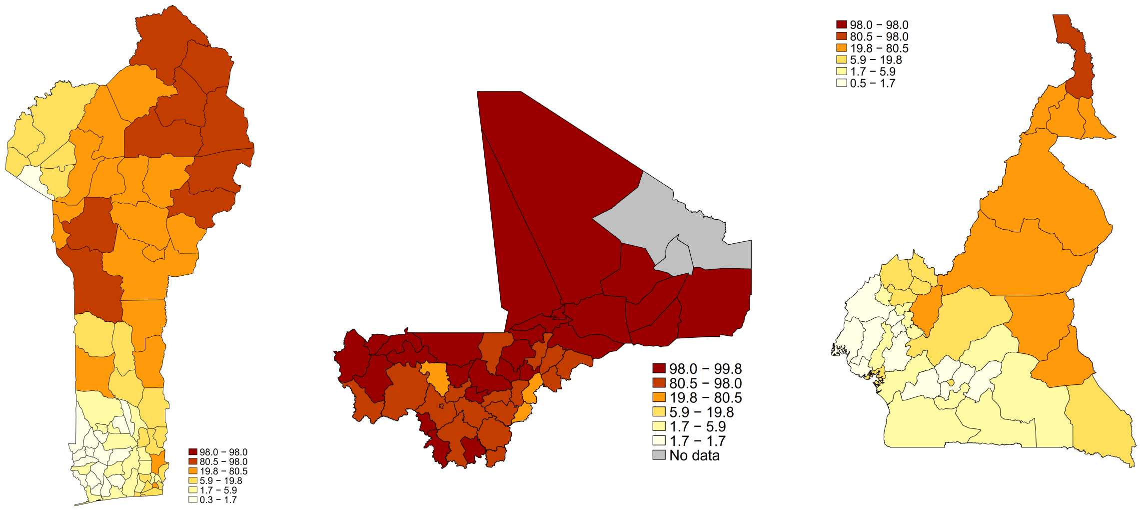
Figure 9 Mapping estimated place effects



(a) Benin (2013) (b) Mali (2009) (c) Cameroon (2005)

Note: These figures map the estimated yearly causal place effects. The left-hand side image shows Benin in 2013, the middle image Mali in 2009, and the right-hand side image Cameroon in 2005. Missing data is the result of dropping localities with too few movers to accurately measure place effects.

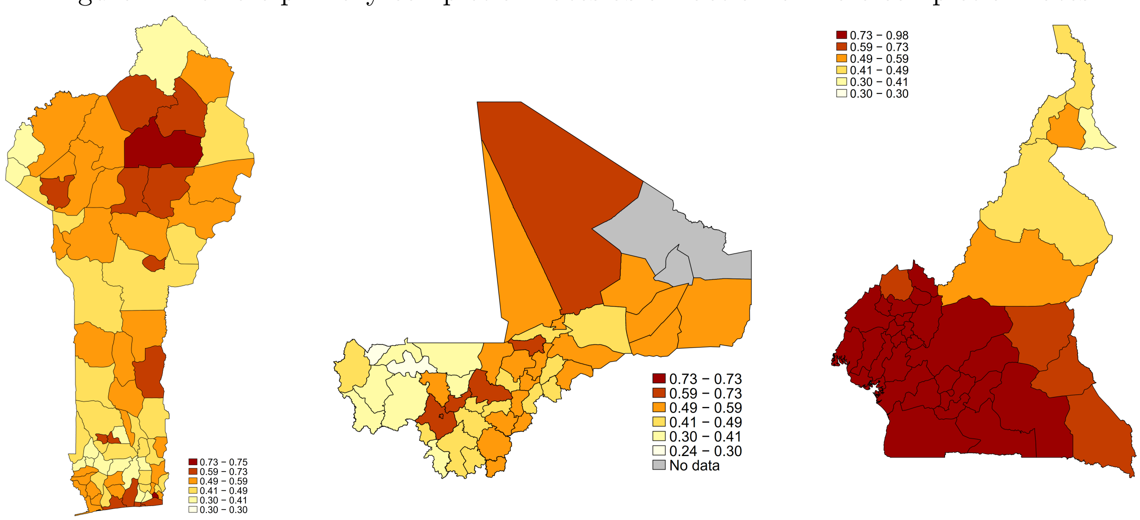
Figure 10 Proportion of the population following Islam



(a) Benin (2013) (b) Mali (2009) (c) Cameroon (2005)

Note: These figures map the observed proportion who report being Muslim in a locality. The left-hand side image shows Benin in 2013, the middle image Mali in 2009, and the right-hand side image Cameroon in 2005.

Figure 11 Female primary completion rates as a fraction of male completion rates

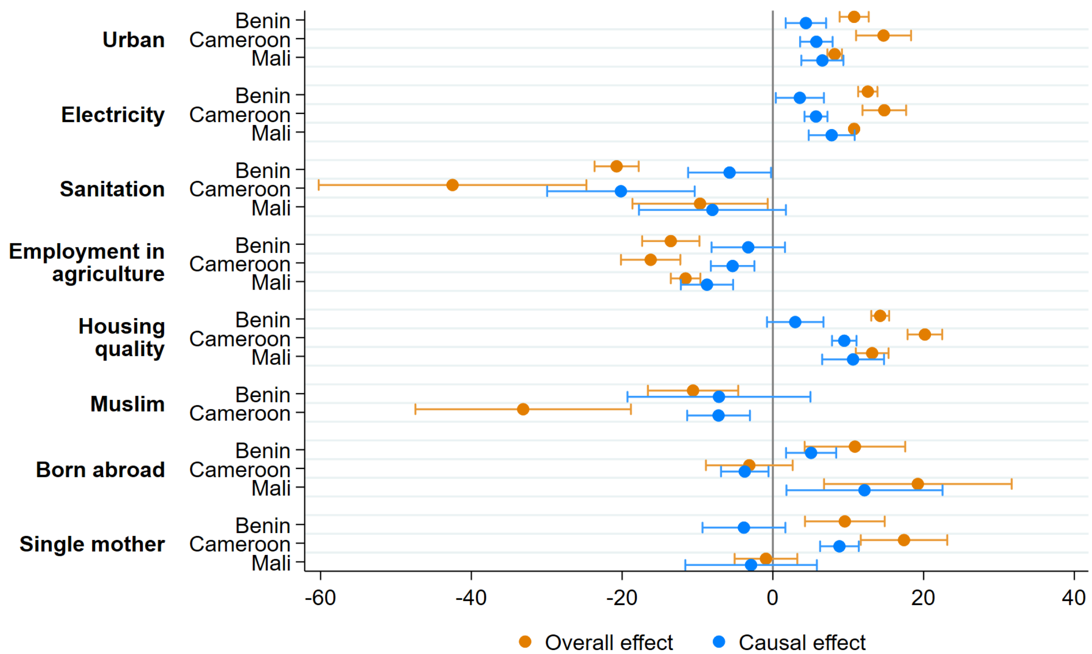


(a) Benin (2013) (b) Mali (2009) (c) Cameroon (2005)

Note: These figures map the observed ratio of reported female to male primary completion in each locality. The left-hand side image shows Benin in 2013, the middle image Mali in 2009, and the right-hand side image Cameroon in 2005.

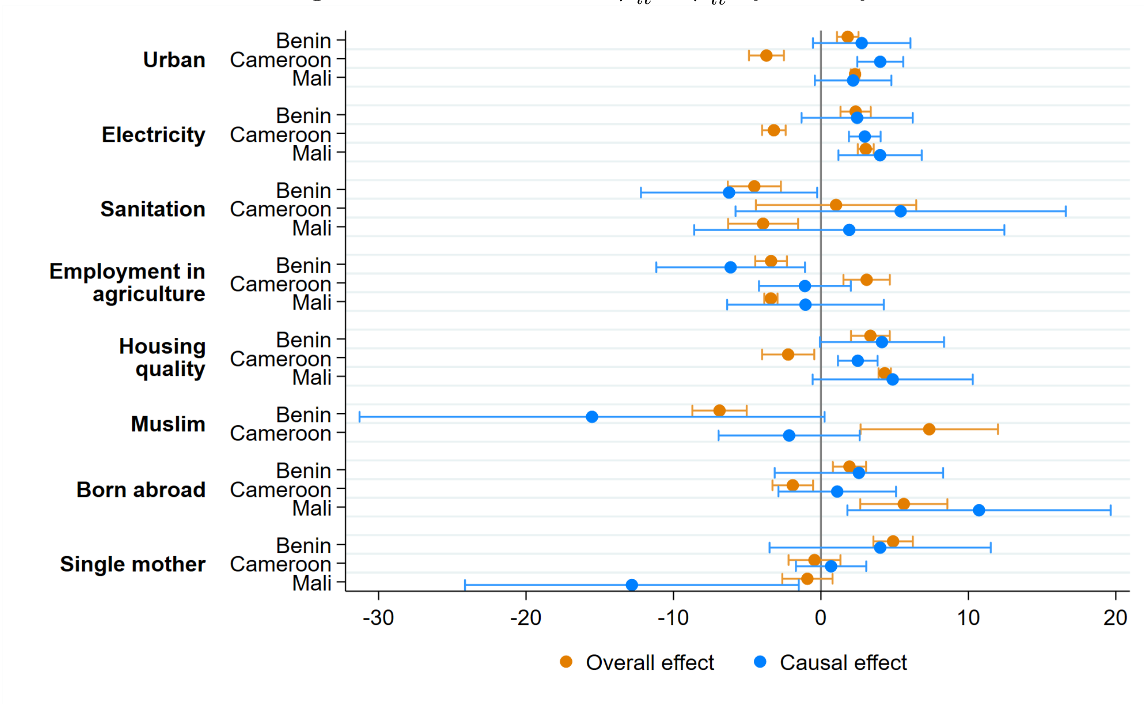
8 Heterogeneity by country analysis

Figure 12 Correlates with overall and causal place effects by country



Note: This figure shows the results from running regressions of observed permanent resident primary completion rates and estimated causal place effects against place characteristics including census fixed effects. Each dot (with corresponding confidence intervals) represents a coefficient from an individual regression with the explanatory variable indicated on the y-axis. In orange, the left-hand side variable is observed permanent residents' primary completion, and in blue it is estimated causal place effects multiplied by 13. Standard errors are clustered at the locality level.

Figure 13 Correlates with $\mu_{lt}^{m} - \mu_{lt}^{f}$ by country



Note: This figure shows results from running regressions of the difference between male and female relative causal place effects on various place characteristics, allowing coefficients to vary by country. Each dot (with corresponding confidence intervals) represents a coefficient from an individual regression with the explanatory variable indicated on the y-axis. Census fixed effects are included and standard errors are clustered at the locality level.

9 Investigation of inequality of opportunity in Cameroon in 2005

Figure 5 in the main text doesn't plot a variance for Cameroon in 2005. This is because there is no significant variation in the causal effect of place on primary completion rates in Cameroon in 2005. This result begs the question of whether Cameroon has suddenly become a much more (spatially) egalitarian society, or if something else is going on. Over time, in all three countries studied, primary completion rates are increasing quite rapidly, leading to less and less variation in primary completion. In Cameroon in 2005, among the individuals who provide identifying variation over 85% completed primary education (as compared to 34% in the entire sample). This means that there simply isn't sufficient variation in outcomes to estimate the variance of individual place effects with any precision. To investigate to what extent place does still matter in Cameroon in 2005, I perform a similar analysis to that in section 3 in the main text but on a sample of 18 to 22-year-olds and looking at secondary completion as the outcome of interest. Table 3 shows the results where I compare the overall causal effect of spending an additional year of childhood in Cameroon in 2005 on primary and secondary school completion. The slope on primary completion, although significantly different from zero, is about half found using the full sample. I find that place also matters for secondary school completion. This indicates that the low variance of place effects for primary school does not suggest that in 2005 where you grow up doesn't matter in Cameroon, but that it matters less for primary school completion as the overall rate of completion approaches one.

Table 3 Causal effect of place on primary and secondary completion in Cameroon in 2005

	(1) Primary	(2) Secondary
1 to 12 slope, primary	-0.0145**	
	(0.00697)	
13 to 18 slope, primary	0.0134	
	(0.0100)	
1 to 12 slope, secondary		-0.0173**
		(0.00602)
13 to 18 slope, secondary		-0.00108
		(0.0206)
Observations	26035	25888

Note: This table shows the results of estimating the linearly parameterized place effects model in Cameroon in 2005 where the outcome of interest is either primary school completion (sample of 14 to 18-year-olds) or secondary school completion (sample of 18 to 22-year-olds).

The above analysis also asks the bigger question: to what extent does the variation discussed in the main text capture true variation in opportunity across space and time, and to what extent does it merely reflects our ability to measure opportunity through variation in primary completion as observed rates increase? Table 4 shows summary statistics of the identifying variation across each census used where the sample has been restricted to that giving identifying variation. This table shows that the standard deviation of primary completion in the identifying sample is large and stable for all samples except Cameroon in 2005 — a lack of identifying variation may not be driving differences found in figure 5 in the main text. Table 4 also shows that primary completion is an important margin in all samples with Cameroon in 2005 being by far the highest, suggesting that it remains associated with opportunity in all other samples.

Table 4 Estimating sample primary education variation

Sample	Mean Primary completion	Sd Primary completion	N
Cameroon 1976	0.61	0.49	32,550
Cameroon 1987	0.72	0.45	61,270
Cameroon 2005	0.86	0.35	157,280
Benin 1992	0.35	0.48	23,950
Benin 2002	0.48	0.50	36,970
Benin 2013	0.72	0.45	41,790
Mali 1998	0.27	0.44	16,630
Mali 2009	0.42	0.49	47,450
Total	0.67	0.47	417,890

Note: This table shows summary statistics of the variation in primary school completion for each census. The sample has been restricted to one-time movers between the ages of 14 and 18 in each case.

References

Raj Chetty and Nathaniel Hendren. The impacts of neighborhoods on intergenerational mobility i: Childhood exposure effects. *The Quarterly Journal of Economics*, 133(3):1107–1162, 2018a.

Raj Chetty and Nathaniel Hendren. The impacts of neighborhoods on intergenerational mobility ii: County-level estimates. *The Quarterly Journal of Economics*, 133(3):1163–1228, 2018b.

Joan Hamory, Marieke Kleemans, Nicholas Y Li, and Edward Miguel. Reevaluating agricultural productivity gaps with longitudinal microdata. *Journal of the European Economic Association*, 19(3):1522–1555, 2021.

David Lagakos. Urban-rural gaps in the developing world: Does internal migration offer opportunities? *Journal of Economic perspectives*, 34(3):174–92, 2020.