The geography of women's opportunity: evidence from Indonesia

César Garro-Marín*

This version: October 7, 2022

Click below to get the latest version
Tables in text   Tables at the end

Abstract

Large and persistent differences exist in women's labor force participation within multiple countries. These persistent differences in employment can arise if where women grow up shapes their work choices. However, they can also arise under endogenous sorting, so that women who want to work move to places where more women work. In this paper, I use rich data from Indonesia to argue that the place women grow up in shapes their participation in the labor market as adults. To do so, I leverage variation coming from women moving across labor markets to estimate the effect on women's labor force participation of spending more time in their birthplace. My strategy is similar to that of Chetty and Hendren (2018) and compares the labor supply choices of women who currently live in the same location, but who emigrated from their birthplace at different ages. My results indicate that birthplace has strong and persistent effects on adult women's labor supply. By the time they turn sixteen, women born in a location at the 75th of female employment will be 4 to 10 p.p. more likely to work than those born in a 25th percentile location. Place is particularly important during the formative period between 9 and 16 years old. These results suggest that between 21 to 45 percent of the current spatial inequality in women's employment is transmitted to the next generation growing up in these locations.

Keywords: gender inequality, local labor markets, place effects

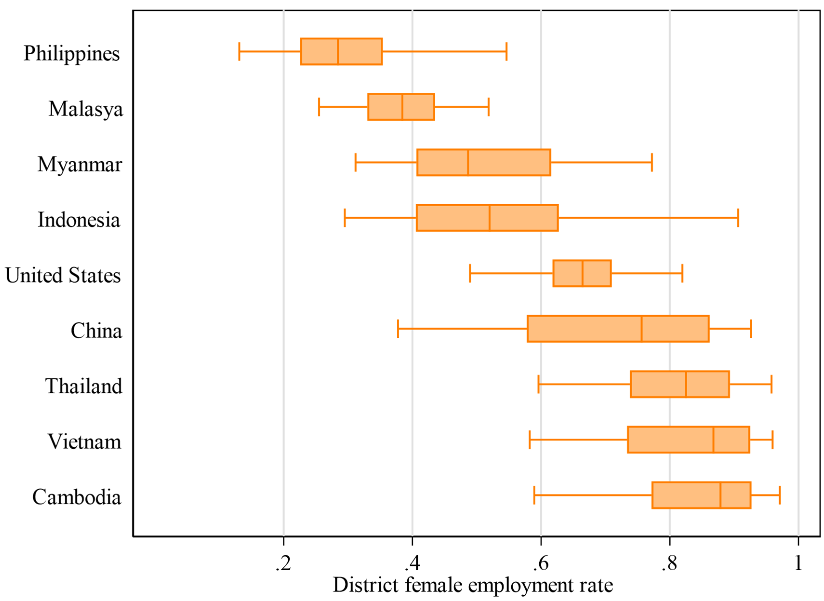
JEL Codes: J16, R19, O18

*Boston University, email: cesarlgm@bu.edu. I am indebted to my patient advisors Daniele Paserman, Kevin Lang, and Linh Tô for their constant support and helpful comments on this project. Special thanks to Bilge Erten, Dilip Mookherjee, Nils Lehr, and Masyhur Hilmy and for enlightening conversations and suggestions. I also thank participants of the BU's Empirical Microeconomics Workshop for helpful feedback. Special thanks to Francesco Benatti for excellent research assistance. All errors are my own.

1 Introduction

There are surprisingly large and persistent differences in female labor force participation (FLFP) rates within countries. In the United States, 76% of women work in the commuting zone of Minneapolis, while only 64% do so in the commuting zone Los Angeles. Other countries feature even more striking differences. In Italy, 59% of women work in the region of Lombardy, but only 31% in Campania. In Indonesia, 56% of women work in the city of Denpasar; but only 43% in the country's capital of Jakarta. These large differences are not exclusive to these three countries. Figure 1 shows that large differences in sub-national employment rates are pervasive in many others. While it is well known that female employment rates vary widely across countries, remarkably little is known about what drives these large disparities within them (Charles et al., 2018).

Figure 1: Female employment rates at the district level for selected countries



Note: The figure shows the distribution of female employment rates for a large subset of Asian countries with geographic data available in IPUMS International. Countries are ordered by median district employment rate. I use the latest available sample from IPUMS International for each country. I aggregate data at the smallest geographical unit available which often corresponds to a district/county, except in the United States where I aggregate data for the US into Commuting Zones as in Autor and Dorn (2013).

This large dispersion in FLFP could be driven by multiple sources. One option is that there are systematic differences in the types of women that locate in different places in these countries. This can happen if, for example, more educated women are more likely to work, and they tend to locate in specific areas of the country. Another possibility is that place shapes women's employment choices. A growing literature in economics argues that current spatial inequality in outcomes such as crime, health, income, and education can have effects on children growing up in these areas (Damm and Dustmann, 2014; Chetty et al., 2016a,b,c; Chetty and Hendren, 2018a,b). Therefore, the current differences in women's labor force participation could have affected the employment choices of women born in these places. This could have important implications. On the one hand, policies encouraging women's incorporation into the labor market should be tailored more locally.

In this paper, I use data from Indonesia to study how women's participation in the labor market is shaped by the place they are brought up. To identify the causal effect of place I leverage variation coming from women moving across labor markets to estimate the effect on women's labor force participation of spending more time in their birthplace. My strategy is similar to that of Chetty and Hendren (2018a) and compares the labor supply choices of women who currently live in the same location, but who emigrated from their birthplace at different ages as children. This allows me to distinguish the causal effect of place from variation driven by differences in women's characteristics whenever women by exploiting variation in the time spent in the birth location.

My results indicate that birthplace has strong and persistent effects on adult women's labor supply. I first show that where women's birthplace is highly predictive of their employment choices. Conditional on living in the same local labor market, women born in high-female employment locations are much more likely to work than those born in places with low-female employment. This strong relationship arises even for women who emigrated young, at ages where women are unlikely to be working. I call this birthplace persistence. Next, I argue that this relationship reflects the causal effect of birthplace on women's employment. To do so, I exploit data on the exact time they left. By comparing birthplace persistence of women who migrated at different ages, I estimate the effect of spending more time in their birthplace. Similar to Chetty and Hendren (2018a) these estimates rely on the assumption that the degree of selection is independent of the age at which women move. For example, this means that I allow for women, or women's families, from high-employment places to be better at choosing their destination than their counterparts born in low-employment places. What I require is that these skill gaps be the same for women who emigrated at 6 years old versus 12 years old. I find considerable support for this assumption in my data.

I find that spending late childhood and early teen years in high-employment locations makes women more likely to work as adults. Staying in a location at the 75th percentile of female employment between 6 and 16 years of age makes women between 4 to 10 percentage points more likely to work than those born in a location at the 25th percentile. These magnitudes mean that between 21 and 45% of the current spatial inequality in women's employment transmits to the next generation of women. In contrast, there are no such effects for men. Therefore, these women-specific place effects can be an important driver of the large and persistent differences in women's labor force participation within countries.

I take advantage of rich Indonesian data detailing people's birthplace and current location at a detailed geographic level. My main analysis sources data from all waves of the Indonesian Family Life Survey (IFLS) These are representative and publicly available datasets that track respondents' birthplace, current location, and migration history across midsize geographies. This level of detail allows studying differences in women's labor supply and the effect of place at a level that is not possible in other countries from standard sources (Bryan and Morten, 2019). Throughout the paper, I identify localities as Indonesian "regencies". There are medium sized administrative geographies akin to counties in the United States. The average regency is approximate twice the size of the US state of Rhode Island and houses eight hundred thousand people.

This paper contributes to three strands of the literature. First, I contribute to the growing research showing that where women grow up and live can influence outcomes such as human capital investment, employment, and fertility choices (Molina and Usui, 2022; Charles et al., 2018; Boelmann et al., 2021). I make two main contributions to this literature. First, I provide evidence of when does women's surrounding environment matter in shaping their choices in adulthood in a large developing country. By applying techniques borrowed from the place effects literature (Chetty and Hendren, 2018a,b; Milsom, 2021), my results suggest that the influence of women's origin place is key during late childhood and early teens. Although previous research has pointed out that women's childhood environment matters for their adult outcomes, this literature is mostly silent on *when* does it matter. Second, my results provide new evidence that where women grow up can matter at a very local level. Previous research emphasizes that differences in norms, culture and other factors across large geographical areas such as states, provinces, or countries can shape women's choices (Charles et al., 2018; Boelmann et al., 2021; Alesina et al., 2013). By exploiting much more disaggregated data, my results suggest these factors can act at a more local level.

Second, this paper also contributes to the literature on place effects. Primarily using evidence from developed countries, this literature shows that where people grow up and live has important implications for intergenerational mobility (Chetty and Hendren, 2018a,b), racial inequality (Chetty et al., 2020), human capital accumulation (Molina and Usui, 2022), and criminal activity (Damm and Dustmann, 2014). I add to this literature by providing new empirical evidence linking women's early environment to their outcomes as adults in a large developing country. In this way, my findings complement existing work showing that spatial inequality is particularly important for women's human capital investment in West Africa (Milsom, 2021) and Japan (Molina and Usui, 2022).

Finally, my paper also adds to the vast research linking gender inequality and culture. This literature provides extensive evidence that gender norms and culture can transmit from parents to children and thus reproduce existing patterns of gender inequality (Fogli and Veldkamp, 2011; Alesina et al., 2013; Fernandez and Fogli, 2009; Charles et al., 2018). My results suggest that the local environment women grow in can have an effect over and above the direct vertical transmission of norms from parents to daughters. This provides suggestive evidence that the transmission of local culture and norms from fellow residents can also be an important factor at reproducing the large gender inequality seen in developing countries. These findings are thus complementary to those of Boelmann et al. (2021), who provide strong evidence that women arriving to places with a stronger female-work culture as adults do internalize these norms.

2 Data

My analysis requires three main pieces of information: (i) detailed data on people's birthplace and place of current residency, (ii) data on employment and migration histories, and (iii) characteristics of the places people live or were born. My main dataset combines data on birthplace, place of residence, employment, and migration coming from the Indonesian Family Life Survey (IFLS). I supplement the IFLS data with place characteristics coming the Indonesian Census and the National Socioeconomic Survey (SUSENAS).

The IFLS is a rich panel survey tracking the information of approximately 34 thousand Indonesians across five survey years: 1993, 1997, 2000, 2007, and 2014. Overall, the IFLS is representative of 83% of the Indonesian population[1] (Hamory et al., 2021). The IFLS has three advantages that make it uniquely suitable to study place effects on female labor supply. First, it records the birthplace and location of residency in small geographic units. The IFLS records people's place of birth and residency at the level of the "regency". These are Indonesian administrative units similar to US counties that allow me to study differences in women's employment between small geographic units[2]. The typical regency is home to approximately eight hundred thousand people and covers an area roughly twice size of the US state of Rhode Island. Second, all people's migration episodes since they were 12 years old, their location at twelve years old, and their work history in all the years since the last survey year. With this information I build a panel dataset tracking migration since birth and the work histories of each individual for up to 26 years, from 1988 to 2014. Finally, the IFLS also contains detailed information on childhood conditions, marriage history, fertility, etc. Thus, I can control for a rich set of potential confounders.

I source place characteristics data from the Indonesian Decennial Censuses and SUSENAS. I use data from the 1980, 1990, 2000 and 2010 Indonesian Censuses available in IPUMS International (Minnesota Population Center, 2020), and data from the 2012, 2013, and 2014 waves of SUSENAS (Badan Pusat Statistik, 2019, 2020). The Censuses and SUSENAS are very similar to each other. Both are large and nationally representative datasets that track the regencies of birth and current residency for all respondents. However, while SUSENAS has smaller samples, it contains a richer

[1]The IFLS originally sampled households from 13 of the 27 provinces that existed in 1993. These provinces account for 83% of the Indonesian population. Appendix figure B.1 shows the geographic area covered by the original IFLS sample. Subsequent waves include data from outside this area whenever members from the original 1993 households moved to these locations. While most of my respondents are located only in these provinces, appendix figure B.2 shows a considerable amount of respondents were born outside these areas.

[2]Datasets available for other countries track geographic information only for states or provinces, which in most cases are either too big or too few to be interesting (Bryan and Morten, 2019)

set of information not available in the Census. For example, data on wage, age at first marriage, fertility, etc., is available only on SUSENAS. I compute all regency characteristics by aggregating these datasets at the regency-level. Whenever possible, I compute these aggregates from the Census, restricting the sample to people aged 18 to 64 years old.

My main measure of women's labor supply is a dummy equal to one if she was employed during the year. This is the variable that I can track most consistently across time in the IFLS. However, as a robustness check, I also show that my results carry through to alternative measures such as being a paid worker, total weekly hours worked, and being a full-time worker.

Throughout the analysis, I associate women's labor supply choices with the characteristics of their birthplace. This exercise requires having geographic units with boundaries that are fixed across time. However, regencies' boundaries changed considerably from decade to decade between 1980 and 2010. Moreover, the creation of new regencies was a common event. Appendix table 11 shows that, just between 2000 and 2010, 154 new regencies were created. I address this issue by using regency aggregates with consistent boundaries between 1970 and 2010. These regency aggregates were built by IPUMS International (Minnesota Population Center, 2020). In total there are 268 consistent-boundary regencies which are just slightly larger than the "original" regencies in the data. From now on, I simply refer to these regency aggregates as regencies.

Table 1: IFLS: summary statistics by gender and migration status

	All (1)	**Women** (2)	**Men** (3)
Age	35.54	35.27	35.85
Married	0.84	0.74	0.97
Attended at least high school	0.37	0.32	0.42
Urban	0.28	0.27	0.28
Muslim	0.89	0.90	0.89
Share left birthplace by age 25	0.31	0.29	0.34
Employed	0.71	0.55	0.89
Type of worker			
Self-employed	0.46	0.40	0.50
Salaried	0.42	0.37	0.46
Unpaid / family worker	0.12	0.23	0.04
Industry of employment			
Agriculture	0.31	0.31	0.32
Services	0.53	0.53	0.53
Manufacturing	0.14	0.16	0.13
Construction	0.05	0.01	0.09
Observations	516,670	276,986	239,684
Number of individuals	37,440	19,074	18,366

Notes: data from IFLS. Urban shows the share of people who report living in a town or a big city. This definition differes from the BPS defined urban classification available in the Indonesian Census and SUSENAS.

In table 1 I summarize the characteristics of my whole dataset, and of women and men separately. Overall, my panel tracks 37,400 individuals, 51% of which are women. They participate in a labor market that is typical of a developing country, with high rates of self-employment (46%) and agricultural work (31%)[3]. There are also large differences by gender in employment rates, worker type and industry of work. The gap in employment is of 34 p.p., which, while large, is in line with the employment gaps in South-East Asia. Moreover, women are twice as likely as men to be unpaid/salary workers. Unpaid workers work or help to earn an income, but who are not paid a wage/salary. Most unpaid workers work in agriculture (82%) and the retail industry (11%) (Minnesota Population Center, 2020). Finally, similar to what has been observed in previous literature, there are gender differences in industry of employment (Olivetti and Petrongolo, 2016, 2014; Blau and Kahn, 2017). In Indonesia, women are overrepresented in manufacturing, while construction is a predominantly male industry.

[3]In comparison, in the United States only 10% of workers are self-employed, and 1% of work in agriculture

Table 2: Indonesia: women's characteristics by migration status

	Stayers (1)	**Migrant** (2)
Age	36.07	35.34
Married	0.75	0.77
Attended at least high school	0.24	0.44
Urban	0.12	0.58
Muslim	0.91	0.87
Share left birthplace by age 25		0.82
Employed	0.57	0.53
Type of worker		
Self-employed	0.43	0.36
Salaried	0.31	0.45
Unpaid / family worker	0.25	0.19
Industry of employment		
Agriculture	0.37	0.20
Services	0.46	0.64
Manufacturing	0.16	0.15
Construction	0.00	0.01
Observations	169,669	68,619
Number of individuals	11,555	6,769

Notes: Stayers are women who *never* left their birthplace. Migrant shows women who live outside their birthplace. Data from IFLS.

My analysis leverages variation that comes from women who emigrated from their birthplace. Table 2 zooms in on this population and compares it to stayers, that is, women who never left their birthplace. There are several differences between migrants and stayers. First, consistent with the findings of previous literature, migrant women are more educated (Hicks et al., 2017). Migrant women are also more likely to be salaried, which suggests that they have access to a less informal labor market. Finally, migrant women display higher shares of urbanicity, lower shares of agricultural employment, but higher shares in services.

Despite what the large shares of urbanicity of women migrants suggest, women's migration is not predominantly from rural to urban areas. In table 3 I follow Bryan and Morten (2019) and classify regencies into urban or rural according to the share of the regency's population who live in areas that BPS (the Indonesian Statistical Agency) labels as urban. I define urban regencies as those whose urban population is above a set cutoff. The table shows the migration rate by type of regency, along with the share of migration episodes that happen between regencies of the same type. Women from urban areas emigrate at similar rates to those in rural areas. Moreover the rural to rural, urban to urban, and urban to rural flows are important. 37% of women born in a rural regency emigrate to another rural regency and 28% of the urban-born move to rural areas.

Table 3: IFLS: women's migration patterns and regency characteristics by urbanicity of regency of origin

	Rural	**Urban**	**All**
	(1)	(2)	(3)
A. Migration			
Migration rate	0.30	0.27	0.28
Within category	0.37	0.72	0.53
B. Characteristics of origin regency			
Women's employment rate			
Average	0.57	0.46	0.52
STD	0.14	0.11	0.14
Number of regencies	135	94	229
Share of IFLS women born in these regencies	0.49	.51	100

Notes: Migration is measured as living outside the regency of birth. In this table, for people observed in more than one regency during their adulthood, I keep the most recent place of residence only. Following Bryan and Morten (2019) I classify regencies as urban if the share of population living in an urban area is above a cutoff. I choose the cutoff to match the urban share at the national level. Data from IFLS and IPUMS International.

3 Four facts about women's labor supply

In this section, I use data from IPUMS International and the 1980-2010 Indonesian Censuses to present four empirical facts on female employment. First, I use data from multiple countries to show that large geographic differences in women's employment rates within countries are pervasive across the world. Next, I zoom in on Indonesia and (i) characterize the large dispersion in female employment across regencies, (ii) document that it is extremely persistent over time, and (iii) show that it is not accounted for by variation in women's demographics or labor market characteristics across regencies. Taken together, these four facts suggest that structural differences could be driving the dispersion in women's labor supply within Indonesia.

3.1 Fact 1: within-country dispersion in women's labor supply is pervasive across countries

In table 4 I use data from IPUMS International to compute summary statistics of women's and men's regional employment rates for several Asian countries and the United States. For each country, I aggregate the employment data at the smallest geographical unit available. For most countries, this corresponds to an administrative region akin to a county or municipality. In the table, I ordered countries from highest to lowest regional dispersion in women's employment.

This table show three important pieces of information. First, Panel A highlights that despite the large differences at the mean, these countries have large regional differences in women's employment rate *within* their borders. For most countries, the gap between the localities at the 75th and 25th percentiles (IQR) is above 15 percentage points (p.p.). A gap of 15 p.p. is large even for high female employment countries such as Vietnam, Cambodia, and Thailand. Even the smaller IQR of 9 p.p. for the United States acquires significance when we consider that this gap is equivalent to the change in the national US female employment rate during the last *thirty-eight years* (1984-2022)[4]. Second, panel A also shows that Indonesia is not an outlier. The large dispersion of female employment in Indonesia is well in line with that of countries such as China, Myanmar and Vietnam.

Third, panel B shows that the large geographic dispersion is mostly exclusive to women's employment rates. This panel displays similar statistics for men's employment rates. In all countries but the US, the dispersion in women's employment rates is substantially larger than men's. In fact, in six out of the nine countries, the dispersion in women's employment *more than doubles* that of men's. In essence, this indicates that men work at high rates everywhere within these countries. Women, however, work at very different rates depending on the locality they live in.

[4]This benchmark is not affected by the Covid-19 drop in women's employment. By 2022, women's employment had recovered to pre-Covid levels.

Table 4: Dispersion in regional employment rates for selected countries

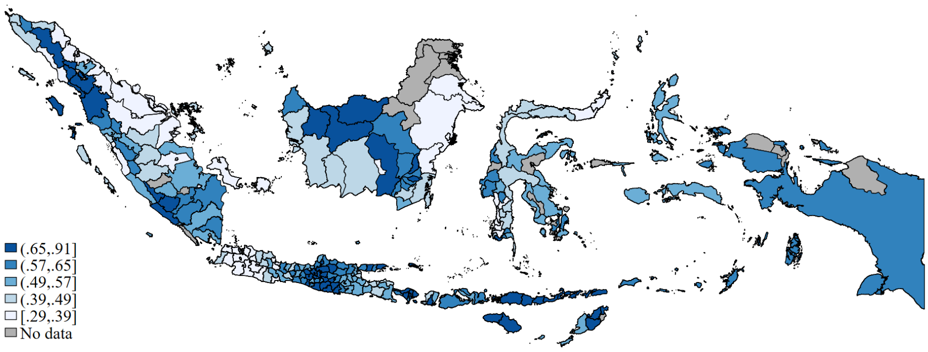
Statistic	**China**	**Indonesia**	**Myanmar**	**Vietnam**	**Cambodia**	**Thailand**	**Philippines**	**Malaysia**	**USA**
A. Women									
IQR	0.28	0.22	0.21	0.19	0.16	0.16	0.13	0.11	0.09
SD	0.17	0.14	0.13	0.12	0.11	0.11	0.10	0.07	0.07
Mean	0.71	0.53	0.51	0.82	0.84	0.81	0.30	0.38	0.67
B. Men									
IQR	0.14	0.05	0.07	0.06	0.08	0.08	0.08	0.06	0.10
SD	0.10	0.04	0.05	0.06	0.05	0.06	0.06	0.04	0.07
Mean	0.85	0.87	0.86	0.90	0.90	0.88	0.82	0.84	0.77
Mean population 18-64	266,748	533,867	83,531	79,146	50,186	58,290	40,423	91,509	202,635
No. districts	2,845	268	362	674	174	670	1,274	133	722

Notes: SD and IQR stand for Standard Deviation and Interquartile Range. The table shows statistics from a large subset of Asian countries in IPUMS International with data available at a small geographic level plus the United States. Columns are ordered from highest to lowest dispersion in women's labor supply. I use the latest available sample from IPUMS International for each country. I aggregate data at the smallest geographical unit available, except for the USA where I use Commuting Zones as in Autor and Dorn (2013). I show the unweighted cross-locality means which –might– differ from the national-level means.

3.2 Fact 2: there is large within-country dispersion in women's employment rates in Indonesia

Figure 2 zooms in on the large dispersion in female employment rates in Indonesia. In this figure I plot women's employment rates across all the 268 regencies in my dataset. Each color groups a quintile of the regencies. Darker blues denote higher employment rates. This map makes evident that women work at very different rates across the country. Women in the top quintile of regencies see employment rates higher than 65% while those living in the bottom quintile see rates below 29%. These low-employment regencies include important population centers, such as Bogor regency and the city of Medan[5]. The map also illustrates that the dispersion in women's employment extends across the whole country and it is not driven by any particular province, island, or group of regencies.

Figure 2: Indonesia: women's employment rate by regency, 2010



Note: The figure shows regency-level employment rates for women aged 18-64. It shows all the 268 regencies with consistent boundaries between 1970 and 2010. Each color groups a fifth of the regencies. The figure uses data from the 2010 Indonesian census from IPUMS international.

3.3 Fact 3: women's employment rates are highly persistent

The large dispersion in women's employment rates could be the result of: (i) temporary economic shocks that depress women's employment in some parts of Indonesia, (ii) measurement error in the employment rates, or (iii) structural differences across regencies correlated with female employment. If the dispersion arises predominantly due to temporary shocks or measurement error, then we should expect very low persistence in the regencies' employment rates across years. This is because any temporary shock should dissipate after a several years and I expect measurement error to be independent across time. In contrast, high cross-year persistence can be taken as evidence that the dispersion in women's employment reflects structural differences across regencies.

[5]Medan is the capital and largest city in the province of North Sumatra. As of 2020, it is also the third most populous city in Indonesia (Brinkhoff, 2022). Bogor is a regency with more than 5 million people. It borders the Jakarta metropolitan area. See their locations in figure B.4 in the appendix.

Table 5: Indonesia: autocorrelation in regency-level men's employment rate, 1980-2010

Regressor	(1)	(2)	(3)	(4)
Female employment 10 years ago	0.80			
	(0.02)			
Female employment 20 years ago		0.72		
		(0.03)		
Female employment 30 years ago			0.70	
			(0.04)	
Same-year male employment				0.51
				(0.04)
Observations	800	534	268	1,071

Notes: The table shows the autocorrelation of regency-level employment rates across different time horizons. It also shows the simultaneous correlation between the employment of both genders. Data from 1980-2010 Indonesian Census taken from IPUMS international. Robust standard errors in parenthesis.

In columns (1) to (3) of table 5 I show estimates of the autocorrelation of the regency-level employment rates across different time horizons. These come from running regressions of the form:

$$e_{rt} = \gamma_{t-j} e_{rt-j} + \varepsilon_{rt} \tag{1}$$

where e_{rt} is the standardized employment rate in regency r at time t.

The autocorrelation estimates suggest the dispersion in women's employment rates reflects structural differences across regencies. The estimates are very high, starting at 80% for ten years, and staying as high as 70% for thirty years. As a benchmark, in column (4) I also show the estimate of the simultaneous correlation with men's employment rates. Note that women's employment rates are *more* correlated with themselves 30 years apart, than with the *same-year* male employment rates.

3.4 Fact 4: dispersion in women's employment rates cannot be accounted by differences in women's demographics alone

The large persistence in the female employment rates suggests that the dispersion in female employment rates reflects structural differences across regencies. These might be differences in the family structure or features of the labor demand across regencies. Motherhood is associated with lower attachment to the labor market by women (Angelov et al., 2016; Kleven et al., 2019). Moreover, differences in the industry mix account for up to 80% of the variation in women's labor supply in developed countries (Olivetti and Petrongolo, 2016). The dispersion in female employment rates might simply reflect differences in the family structure or the industry mix across regencies.

In table 6 I test whether permanent differences in the industry mix or women's demographics can account for most of the dispersion in employment across regencies. This table shows the R^2 from regressions of employment rates on a series of controls. They include the share of people married, the share with small children, along with measures of the age structure, the education level by gender, and the industry mix. I run the regressions separately by gender and I stack data from all 1980-2010 censuses. Additionally, I include year fixed-effects to capture national trends in employment.

Table 6: Indonesia: share of employment rate dispersion accounted for observed regency characteristics, 1980-2010

	Women's					**Men's**				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
R^2	0.13	0.26	0.30	0.31	0.47	0.01	0.41	0.60	0.69	0.79
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Age structure		✓	✓	✓	✓		✓	✓	✓	✓
Women's education level			✓	✓	✓					
Men's education level								✓	✓	✓
Share married				✓	✓				✓	✓
With child under 5				✓	✓				✓	✓
Industry shares					✓					✓
N	804	804	804	804	804	804	804	804	804	804

Notes: The table reports the R^2 of a regression of regency employment rates on regency-level aggregates. Age structure controls are shares aged 30-49 and 50-64. Education level measures are shares who attended at most middle school, high school, and college. When indicated, the regressions include 1-digit industry shares. Sample restricted to people aged 18-64. Data from IPUMS International.

Table 6 shows that differences in women's demographics or the industry mix account for at most a moderate share of the dispersion in female employment across regencies. Column (4) shows that controlling for women's education level, and the regency's family and age structure accounts for only a third of the dispersion in employment rates. In column (5) I add a full set of 1 digit industry shares as controls. Adding the industry mix increases the R^2 by 16 p.p. While this increase in explanatory power indicates that the regency industry mix is an important factor behind the variation of women's employment rates; collectively, all regressors I consider leave 53% of the dispersion in women's employment rates unaccounted for. In contrast, when I perform a similar exercise for men's employment rate, column (10) shows that the *same* factors leave only 21% percent of the men's dispersion unaccounted for. In all, this table suggests that the dispersion in women's employment reflects structural differences across regencies that *predominantly affect women's labor supply.*

4 Empirical strategy and results

In this section, I present evidence of the large and persistent effects of birthplace on women's labor supply using data from women migrants in Indonesia. I start by illustrating how I identify these effects using information from women who currently live outside their birthplace. Next, I present the empirical evidence in three steps. First, I show that conditional on the current place of residence, birthplace is highly predictive of women's labor supply in adulthood. This persistence can reflect both the causal effect of birthplace or a spurious correlation driven by women's unobserved characteristics. Next, I build on this result and show that birthplace also has high predictive power for women who left their birthplace before they turned 18, a sample for which migration driven by current employment opportunities is less prominent. Finally, using a strategy similar to (Chetty and Hendren, 2018a), I show that, for these early women migrants, birthplace has stronger predictive power the longer they stayed in their birthplace. These estimates rely on the assumption that the degree of selection is independent of the age these women left their birthplace. I present evidence providing considerable support for this assumption.

4.1 Place and women's labor supply: the identification challenge

The place of residence can have both direct and indirect effects on women's labor supply. Direct effects are those that affect the labor supply of all the current female residents. There is considerable empirical evidence documenting these effects. These might arise, for example, from factors such as the levels of childcare availability (Compton and Pollak, 2014), commuting costs (Le Barbanchon et al., 2021; Farre and Ortega, 2021), the industry makeup (Olivetti and Petrongolo, 2014), or the level of sexual discrimination in the local labor market (Charles et al., 2018). Differences across localities in any of these factors will cause geographic differences in women's labor supply across localities. However, place can also affect women indirectly by affecting their preferences and the skills they acquire. Women born and brought up in locations where many women work can internalize these norms and thus be more likely to work as adults (Charles et al., 2018; Boelmann et al., 2021). Moreover, environments with high female employment could make women more likely to invest in acquiring the skills they need to participate in the labor market (Molina and Usui, 2022). These permanent indirect effects will create differences in labor supply across women born in different locations *irrespective* of where they currently reside. Evidence on indirect effects is much more scarce in the literature (Charles et al., 2018).

The selection problem

In this paper, my main interest lies in determining what women's labor supply would be if, conditional on the current place of residence, she was born in an area where more women work. This counterfactual exercise keeps the woman, her family, and her place of residence fixed and varies only her childhood experience. To answer this question, I study the labor supply of women residing outside their birthplace. Because for these women the place of residence is different from their birthplace, I can separate the indirect effects from the direct effects of place. More formally, let us consider the following model for women's probability of employment e_i,

$$e_{it} = \delta_c + \boldsymbol{\sigma} p_b + \eta_{it} \tag{2}$$

in this model women's employment choices depend on three main factors. First, a place-of-residence fixed effect δ_c that captures all the direct effects of location c on female labor supply. These might include commuting costs, childcare availability, and gender discrimination, among others. Second, the birthplace female employment p_b captures the causal effect of growing up in a location where p_b percent of the women work. Finally, the error term η_{it} captures all other factors making some women migrants more likely to work than others.

Model (2) follows closely the tradition brought forth by the "epidemiological" approach literature (Fernández and Fogli, 2006). Women's birthplace could have multiple impacts on women's behavior as adults. Including the prevailing female employment rates as the main regressor in equation (2) relies on the idea that these rates capture the place-driven factors key to determining women's employment choices. This specification also facilitates testing whether alternative channels are driving the relationship with the birthplace employment rates (Fernández, 2007).

In model (2) $\boldsymbol{\sigma}$ is the parameter capturing the birthplace effects. It gives the counterfactual increase in employment of women had they been born in a place with 1 p.p. higher female employment rate. In the ideal, but clearly unfeasible experiment, I would reassign women's birthplace randomly while keeping the current residency fixed. Random assignment would guarantee that women's birthplace is uncorrelated with the error term. Thus an OLS regression of (2) gives a consistent estimate of $\boldsymbol{\sigma}$. In observational data, however, it is likely that the unobserved factors imbued in the error term are correlated with birthplace labor supply. Therefore, the OLS will conflate the causal effects of birthplace with selection on unobservables:

$$\begin{aligned} \text{plim } \hat{\boldsymbol{\sigma}} &= \boldsymbol{\sigma} + \frac{\text{cov}(\tilde{p}_b, \tilde{\eta}_{it})}{\text{var}(\tilde{p}_b)} \\ &= \boldsymbol{\sigma} + \boldsymbol{\gamma} \end{aligned} \tag{3}$$

where tilde accents denote variables residualized from regency fixed effects (Angrist and Pischke, 2009). Expression (3) shows that the OLS coefficients $\boldsymbol{b}$ reflects two factors. First, the causal effect of birthplace $\boldsymbol{\sigma}$, but also differences in unobservable characteristics across women from different origins $\boldsymbol{\gamma}$. The key identification challenge is separating the selection term $\boldsymbol{\gamma}$ from the birthplace effect $\boldsymbol{\sigma}$.

The selection term $\boldsymbol{\gamma}$ makes explicit that even in the absence of a causal effect, my birthplace could be capturing characteristics about me or my family that are relevant for my work decision.

I will be arguing later that the causal effect of place is positive ($\boldsymbol{\sigma} > 0$). That is, being born in a place where more women work, makes you more likely to work. In these circumstances, I will be more concerned with omitted variable –or selection– bias making women form high-employment birthplaces more likely to work than their low-employment counterparts. For example, previous research shows that daughters from working mothers are more likely to work (Fernández, 2007). Even in the absence of a causal effect, a positive $\hat{\boldsymbol{\sigma}}$ could simply be reflecting that, in places where more women work, girls are more likely to be brought up by a working mother.

Using emigration age data to identify causal effects

Under additional assumptions, data on the age of emigration allows me to distinguish selection from the causal effect of place. The argument is similar to that of Chetty and Hendren (2018a). I start by assuming that place effects are stronger the longer women stay there. Thus, the employment choice for women who emigrated at age a is determined as:

$$e_{it} = \delta_c + \lambda_a + \boldsymbol{\sigma}_a p_b + \eta_{it} \tag{4}$$

Here $\boldsymbol{\sigma}_a$ captures the cumulative effect of birthplace up to age a[6]. The age of emigration fixed-effects λ_a absorb differences in labor force participation across women who emigrated at different ages. The causal effect of staying in the birthplace at age a is $\boldsymbol{\pi}_a = \boldsymbol{\sigma}_a - \boldsymbol{\sigma}_{a-1}$.

By an argument analogous to that in expression (3) the OLS estimates will conflate the causal effects of birthplace $\boldsymbol{\sigma}_a$ with the selection on unobservables for women migrating at age a $\boldsymbol{\gamma_a}$[7]:

$$\text{plim } \hat{\boldsymbol{\sigma}}_a = \boldsymbol{\sigma}_a + \boldsymbol{\gamma}_a \tag{5}$$

Identification assumption (constant selection)

Selection on unobservables does not depend on the age of emigration, that is $\boldsymbol{\gamma}_a = \boldsymbol{k}$

This assumption essentially requires that, conditional on the location and age of emigration fixed effects, the relationship between the birthplace employment rate and the error term is the same for women migrating at different ages. For example, I generally do not observe whether a woman's mother worked and thus it is in the error term. The mother's labor force participation will be correlated with both the daughter's employment decision and the birthplace employment rate. This *does not* necessarily violate the identification assumption. Constant selection simply requires

[6]The causal effect $\boldsymbol{\sigma}$ in the previous subsection can be interpreted as a weighted average of causal effects by age of emigration.

[7]You can find the full derivation of this expression in appendix section A. I defined γ_a as the a-th element in the vector plim $\left[(\widetilde{P}'\widetilde{P})^{-1}\widetilde{P}'\tilde{\boldsymbol{\eta}}\right]$. Here, P is the matrix containing the interaction between the age of emigration dummies and the birthplace female employment rates. $\boldsymbol{\eta}$ is the vector of error terms. The tildes are variable residualized from current location and age of emigration dummies.

that the relationship between the birthplace employment rate and the mother's labor force status to be the same for women who emigrated at different ages. This assumption would be violated if, for example, girls with working mothers are more likely to stay emigrate later.

Under the constant selection assumption, I can isolate the birthplace causal effect from the selection. By subtracting the OLS estimates across different emigration ages the constant selection term γ goes away, leaving only the causal effects:

$$\begin{aligned} \text{plim } \hat{\boldsymbol{\sigma}}_{\boldsymbol{a}} - \hat{\boldsymbol{\sigma}}_{\boldsymbol{a-1}} &= \boldsymbol{\sigma}_a - \boldsymbol{\sigma}_{a-1} \\ &= \boldsymbol{\pi}_a \end{aligned} \tag{6}$$

this expression also shows that identification does not necessarily require that selection is constant across *all* emigration ages. If instead selection is constant only within some age ranges, I can still identify the effects within those ranges. For example, suppose there is reason to believe selection for women who emigrated between 0 to 6 years old is different than for those who emigrated between the ages of 7 to 15. If constant selection holds *within* these ranges, I can still identify the place effects within the 0 to 6 and 7 to 15 ranges. In section 4.4 I present estimates from the birthplace effects along with evidence that women emigrating at different ages are similar to each other along multiple dimensions.

4.2 Birthplace is highly predictive of women's labor supply

I start by comparing the labor supply of women who *live in the same location*, but who were born in different regencies. I do this by regressing a dummy equal to one if the person is employed at year t (e_{it}) on: regency of current residency fixed-effects (δ_c), year fixed-effects (t), women's employment rate in her regency of birth (p_b)[8], and a set of individual and regency-level controls X_{it}. These controls might include age, religion, education, number of books at home when growing up, among others.

$$e_{it} = \delta_c + \theta_t + \boldsymbol{b}p_b + X_{it}\kappa + \varepsilon_{it} \tag{7}$$

In this regression, $\boldsymbol{b}$ is the parameter of interest. It captures the *association* between women's labor supply and the prevailing female employment rate in their birthplace. For now, I will refer to $\boldsymbol{b}$ as the birthplace persistence coefficient. Because the model includes regency of residency and year fixed-effects, $\boldsymbol{b}$ is –primarily– identified out of differences in labor supply of women who live in the same regency in the same year, but who were born in different localities. In this model δ_c absorbs differences in women labor supply driven by permanent differences in the localities of residency, such as: average wages, industry-mix, healthcare availability, etc., while θ_t captures national secu-

[8] I take the birthplace female employment rate from the 2010 census. However, because the employment rates are so persistent across time –see table 5–, that sourcing these rates from other years produces similar results. See appendix table 15 for more details.

lar trends in women labor supply.

I call the slope of the birthplace employment rate $\boldsymbol{b}$ –rather than $\boldsymbol{\sigma}$– to emphasize that it generally differs from the causal effect discussed in section 4.1. A positive $\boldsymbol{b}$ could reflect differences in factors completely unrelated to the birthplace which make women from high-female employment regencies more likely to work than their counterparts from low-employment regencies. For example, women from high-employment locations are more likely to have working mothers. Alternatively, if working women migrate only when they get a job is likely at the destination, even conditioning on the destination, women from high-employment would be more likely to work than their low-employment counterparts. In the next subsections, I address these challenges using alternative samples and specifications.

Table 7: Indonesia: estimates of birthplace predictive power on women's employment (***b***)

	Women					Men		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Women's employment rate at birthplace (p_b)	0.37***	0.38***	0.35***	0.36***	0.01	0.04	0.05*	0.04
	(0.04)	(0.04)	(0.05)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Mean employment rate	0.54	0.54	0.54	0.54	0.90	0.90	0.90	0.90
Implied IQR gap	0.08	0.08	0.08	0.08	0.00	0.01	0.01	0.01
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Regency FE	✓	✓	✓	✓	✓	✓	✓	✓
Age		✓	✓	✓		✓	✓	✓
Religion			✓	✓			✓	✓
Education				✓				✓
Observations	64,727	64,727	64,727	64,727	60,126	60,126	60,126	60,126
N individuals	6,133	6,133	6,133	6,133	6,293	6,293	6,293	6,293
R^2	0.10	0.12	0.12	0.14	0.05	0.17	0.17	0.18

Notes: Uses data from IFLS. Sample restricted to people residing outside their birthplace. Implied IQR gap shows the implied employment gap between someone born at a regency at the 75th percentile of employment rate and someone born at the 25th percentile. The IQR of the female employment rate across regencies is of 22 percentage points. Standard errors clustered by regency of origin. When indicated, the regressions control for a quadratic polynomial in age, and fixed-effects for seven religion and for education categories. Standard errors clustered by regency of origin.

Table 7 shows estimates of the birthplace persistence coefficient $\boldsymbol{b}$. Column (1) shows results from a baseline specification that includes year and regency of residency fixed-effects only. The coefficient of 0.37 implies that birthplace is highly predictive of women's employment. To see how large this coefficient is, let us consider two women: Freida and Amanda. Freida was born in the city of Probolinggo in East Java, which has a female employment rate of 40%. In contrast, Amanda was born in the regency of Sukoharto in Central Java, which has a female employment rate of 62%. These employment rates place these regencies at approximately the 75th and the 25th percentiles of female employment. The estimate implies that Freida is 8 percentage points less likely to work than Amanda. This is a difference of 15% relative to the employment rate of the average Indonesian woman.

A natural concern with the estimate in column (1) is that it could be primarily reflecting differences in factors, such as education, between women from different origins. Columns (2) to (4) show that this is not the case. Controlling for age, religion, education, and childhood environment barely modifies persistence estimates. That said the estimates in table 7 could still reflect differences in other factors I cannot control for. I address these concerns in the next subsections.

Another question is whether these estimates capture patterns that are specific to women. The additional results in columns (5) to (8) in table 7 show that this is the case. These estimates come from regressions where I relate men's employment in adulthood to their birthplace's *female employment rate*. Remarkably, all these estimates are close to zero and they are precise enough to rule out any large persistence. Moreover, these estimates are so small that they imply little variation in men's employment rates across regencies. For example, the estimate in column (8) implies an IQR gap of just 1 p.p. Given the large employment rates of men, such a gap is economically negligible.

4.3 There is large persistence for those who migrated young

The strong persistence of birthplace in women's labor supply could reflect the fact that working and non-working women have different migration patterns. If working women migrate only when a job is likely at the destination, even conditioning in the current location, women from high-employment regencies would be more likely to work than their low-employment counterparts. This correlation would arise even in the absence of any birthplace causal effect.

Table 8: Indonesia: estimates of birthplace persistence on labor supply for early migrants (***b***)

	Women		**Men**	
	Baseline	**Young**	**Baseline**	**Young**
	(1)	(2)	(3)	(4)
Women's employment rate at birthplace (p_o)	0.36***	0.39***	0.04	0.08*
	(0.04)	(0.06)	(0.03)	(0.04)
Mean	0.54	0.52	0.90	0.86
Implied IQR gap	0.08	0.09	0.01	0.02
Age at emigration	All	< 19	All	< 19
Observations	64,727	27,977	60,119	23,016
N individuals	6,133	2,629	6,291	2,389
N movers	6,133	2,629	6,291	2,389
R^2	0.14	0.16	0.18	0.25

Notes: Uses data from IFLS. Columns (2) and (4) restrict the sample to people who left their birthplace before they turned 19. Implied IQR gap shows the implied employment gap between someone born at a regency at the 75th percentile of employment rate and someone born at the 25th percentile.. All regressions controls for year, regency of residency, religion, and education FE, and a quadratic polynomial on age. Standard errors clustered by the regency of birth.

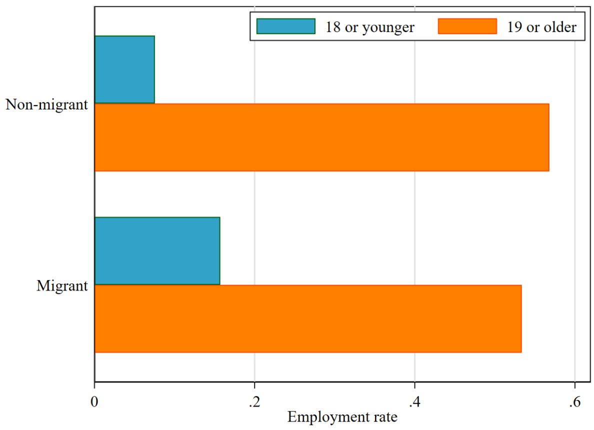
If my results were driven by unobservable differences between working and non-working women, then there should be little birthplace persistence for women who emigrated before they entered the labor market. I test this by restricting the sample to women who emigrated before they turned 19. These women represent 42% of my sample and, as figure 3 illustrates, they were unlikely to be working at the time they emigrated.

In column (2) in table 8, I show the birthplace persistence estimates for these emigres. For comparison, the table also shows estimates for the full sample –Baseline–. Even though these women were unlikely to be working at the time they emigrated, I obtain similar estimates to those of the full sample. Thus unobservable differences between working and non-working women are unlikely to be driving these results. Table 8 also shows estimates of a similar exercise for men. Again, there is little evidence of persistence for men.

4.4 The birthplace persistence is stronger the longer you stay

The strong birthplace persistence in women's employment could still reflect unobservable differences between women from different origins. Here, I address this concern by exploiting differences in the timing of migration to argue that this persistence reflects the causal effect of women's birthplace. To do so, I augment expression (7) by (i) allowing the coefficient on female employment rate to vary by the emigration age ($\boldsymbol{b_a}$), and (ii) adding age of emigration fixed-effects (λ_a).

Figure 3: IFLS: women's employment rates by age and migration status



Note: The figure shows women's employment rates by age and migration status. I define as migrants women who reside outside their birthplace. Non-migrants are those who *never* left their birthplace. Data from IFLS.

As I discuss in section 4.1, I can decompose the OLS estimates of age specific-slopes into a cumulative causal effect σ_a, and a selection term γ:

$$\boldsymbol{b}_a \quad = \quad \boldsymbol{\sigma}_a + \boldsymbol{\gamma}$$

under the assumption that selection is independent of the age of emigration, I can identify the causal effect of place at any given age (π_a) by subtracting the persistence coefficients across emigration ages:

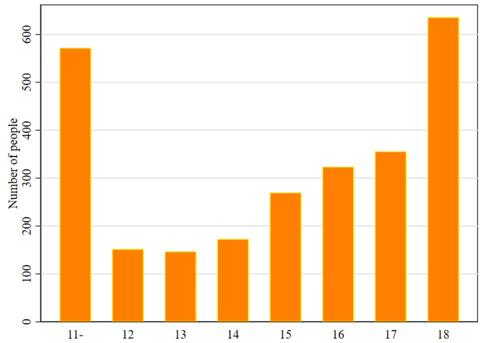
$$\boldsymbol{\pi}_a \quad = \quad \boldsymbol{b}_{a+1} - \boldsymbol{b}_a$$

To estimate this model I leverage age of emigration data from the IFLS. The structure of the survey questionnaire and the sample size imposes several restrictions. First, for all migrants who emigrated before they turned 12, I do not know the exact age at which they left their birthplace. The IFLS only tracks the age at migration for episodes happening after 12 years old. Second, the relatively small sample size in the IFLS forces me to group the age of emigration into bins. This is because the age-specific persistence coefficients are identified out of differences in employment rates between women who (i) live in the same regency, but who (ii) left their birthplace at different ages. Making these comparisons requires within-regency variation in the age of emigration. However, figure 4 shows that the number of women migrating at any given age is small relative to the number of regencies[9]. Therefore, in the estimation I group women into five age of emigration brackets: 11 or less, 12-14, 15-16, 17, and 18 years old. Appendix figure B.5 shows that this group-

[9]On average, only 185 women emigrated in each age between 12 and 16 years old. Because there are 268 regencies, in many places there will be limited within-regency variation to identify different coefficients for each age cell

ing approximately balances the number of women migrants across the age bins.

Figure 4: Indonesia: number of women by age they left their birthplace



Note: Sample restricted to women who outmigrated before they turned 19 years old. Data from the IFLS.

Longer stay does make you more likely to work

Table 9 shows estimates of the birthplace persistence by age of emigration $\boldsymbol{b_a}$. My sample remains restricted to people who left their birthplace before they turned 19 years old. All regressions in the table control for education and religion fixed-effects, along with a quadratic polynomial on age. In column (1) I reproduce the baseline estimates for these women, while Column (2) allows the coefficient on birthplace female employment to vary by age bracket. These coefficients show a striking pattern in birthplace persistence: women born in high-female employment locations are more likely to work as adults the longer they stay in these there.

Table 9: Indonesia: estimates of birthplace persistence ($\boldsymbol{b_a}$) by age of emigration

	Women		**Men**	
	(1)	(2)	(3)	(4)
p_b	0.387		0.082	
	(0.058)		(0.042)	
Age of emigration interactions				
11- $\times$ p_b		0.087		0.109
		(0.113)		(0.082)
12-14 $\times$ p_b		0.217		0.103
		(0.145)		(0.065)
15-16 $\times$ p_b		0.533		0.047
		(0.112)		(0.078)
17 $\times$ p_b		0.572		0.201
		(0.126)		(0.101)
18 $\times$ p_b		0.545		-0.042
		(0.134)		(0.076)
Year FE	✓	✓	✓	✓
Current regency FE	✓	✓	✓	✓
Age of migration FE		✓		✓
Observations	27,977	27,977	23,014	23,014
No. individuals	2,629	2,629	2,389	2,389
No. migrants	2,629	2,629	2,389	2,389
r2	0.16	0.16	0.25	0.25

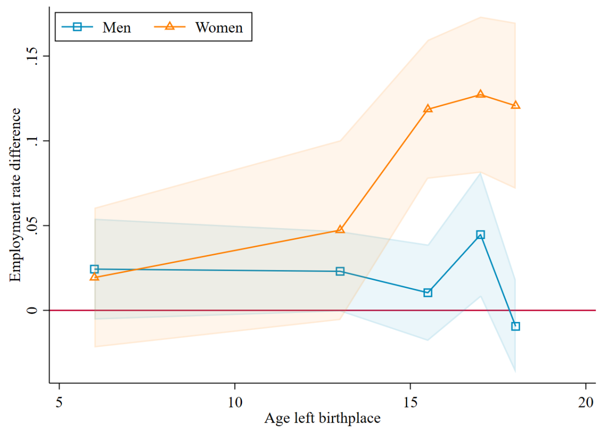
Notes: Uses data from IFLS. Table restricts the sample to people residing outside their birthplace with known age of emigration. All regressions controls for religion, and education FE; and a quadratic polynomial on age.

The results in table 9 suggest that place is particularly important during women's childhood and teenage years. Under the constant selection assumption, the slope estimate for the earliest emigres sizes the selection term, while slope differences provide estimates for the causal effect at specific ages. Therefore, the 11 or less slope provides a benchmark for the selection term. Although I do not know the exact age they emigrated, data from the Intercensal Survey suggests Indonesian emigrate at approximately constant rates in their first 11 years[10]. Thus, the average woman in this bracket stayed only 6 years in their birthplace. From 12 to 16 years the birthplace slopes rise sharply up to 0.53 p.p. to then stay roughly constant thereafter. This points to a positive causal effect of women's birthplace on their adult labor supply. Spending late childhood and early teens in a high-employment regency increases the likelihood women work as adults. There is little evidence of additional place effects after 16 years old. Moreover, these effects are unique to women. In

[10]See figure B.6 in the appendix.

column (4) I perform a similar exercise for men. These slopes do not display the consistent pattern women's slopes have. If anything, these results suggest that men's slopes are driven by selection only.

Figure 5: Indonesia: implied gap IQR in employment rate by age of emigration



Note: The figure shows estimates of $\boldsymbol{b_a}\times$ regency-level IQR in female employment (22 p.p.). This is the implied gap in employment at different ages of out-migration between a person born in a regency at the 75th percentile in female employment, and another one born in a regency at the 25th percentile of female employment. Series for women use $\boldsymbol{\sigma}_a$ estimates in column (2) of table 9. Series for men use $\boldsymbol{\sigma}_a$ estimates from column (4) of table 9. Point estimates are placed at the mean point of the respective age interval. Shaded areas show 90% confidence intervals. The figure uses data from IFLS.

The estimates in table 9 imply that place effects are important drivers of the geographic differences in women's labor supply. I illustrate this in figure 5, where I show the implied gap in employment ages between two women: one born in a regency at 75th percentile of employment and the other at the 25th percentile[11]. I graphed the estimates at the midpoint of each age bracket. If these women had emigrated at 6 years old, there would be a gap of 2 p.p. in their labor supply. This gap is not causal and is fully driven by selection. However, if they left at 13 and 15, the gap would widen to 5 and 12 p.p. respectively. This widening is driven by the causal effect. In all, these magnitudes are very large. They imply that the existing 22 p.p. gap in female employment translates into a gap of 12 p.p. Thus, approximately 45% of the existing inequality in female labor supply is transmitted to the next generation of women growing up in these locations.

The data supports the constant selection assumption

The causal interpretation of the birthplace persistence coefficients hinges on the strong assumption that selection is independent of emigration age. More precisely, conditioning on the current location and other controls, I require that the relationship between women's unobserved characteristics and the birthplace female employment rate be the same for women who emigrated at different ages. Below I provide results showing that selection along several observable dimensions is constant

[11] These two regencies have a gap of 22 p.p. in the female employment rate.

across emigration age. This suggests that the identification assumption is likely to hold in my data.

One can think of the identification assumption as an analog of the parallel trends in Difference in Differences. I expect women coming from high and low-employment regencies to be different from each other. This is not an issue. However, even in the absence of a birthplace causal effect, if there are factors correlated with female employment that change differently across emigration ages for these two groups of women, I would mistakenly assign this variation to the causal effect. In other words, the lack of parallel trends would lead me to find a causal effect where there is none.

Figure 6 shows that women emigrating at different ages display similar selection patterns on (i) reasons for emigrating, (ii) proxies of parental wealth, (iii) number of siblings, and (iv) the characteristics of the destination regency. This figure shows estimates of the slopes on the interaction between age of emigration and birthplace FLFP β_a in regression of the form:

$$y_i = \lambda_a + \beta p_b + \sum_a \beta_a 1_a \times p_b + X_i \kappa + \varepsilon_{it} \tag{8}$$

where y_i is a woman characteristic, and 1_a is an age of emigration dummy. For all regressions I chose the latest age –18– as the base category, thus all the β_a are readily interpreted as the slope difference relative to women who emigrated at 18. *Under constant selection all the interactions must be zero.*

Panel (a) shows estimates where I use indicators of the reason for emigrating as outcomes. This information is available in the IFLS only for people who emigrated at 12 years old or older, thus the figure does not show estimates for the youngest emigre cohort. Although imprecise, these estimates suggest constant selection by emigration motive. For all motives I cannot reject that all interactions are jointly zero[12].

Panels (b) and (c) of figure 6 show estimates for proxies of parental wealth and measures of family size. IFLS respondents provide retrospective information about their household during their childhood. In panel (b) I use as outcomes whether her father had formal employment, an indicator of having more than 11 books at home, and the number of people per room in the household[13]. In all cases, I cannot reject that all the interactions are jointly zero. Panel (c) performs a similar exercise with the number of siblings. Previous research argues that the number of siblings is an important determinant of parental investment's in women's education. Although there seems to be a difference in the number of brothers for 17 year old emigres, reassuringly I cannot reject that all estimates are jointly zero at conventional significance levels.

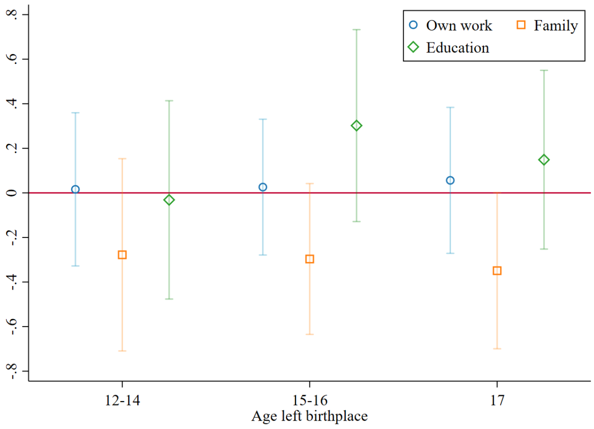
[12]See point estimates and tests of hypothesis in table 16 in the appendix.

[13]Only 21% of the women in this sample declare having 11 or more books when they were 12 years old. I define a respondent's parent as formally if he was employed as a private or government worker. Only 26% of the women in this sample report having a formally employed father when they were 12 years old.

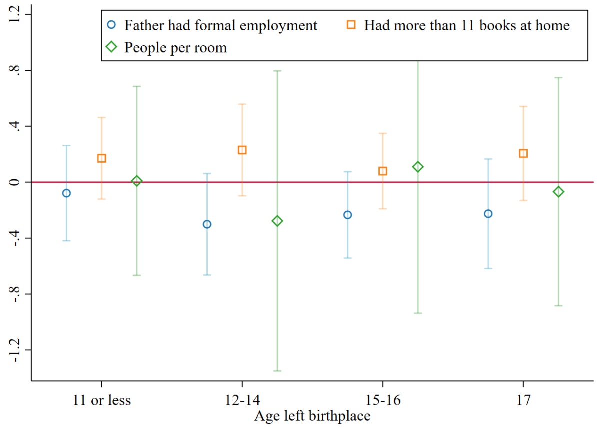
Finally, in panel (c) I explore whether women's destination changes by emigration age. Estimates in this where I use destination regency characteristics as outcomes. Because my birthplace persistence regressions control for the current location fixed-effects, selection due to the destination is less of a concern. It is reassuring, however, that I cannot reject that all the interactions are jointly zero. Thus, I do not see evidence that emigres' age changes the destination they choose.

Figure 6: Indonesia: women and selection by age of emigration in the IFLS

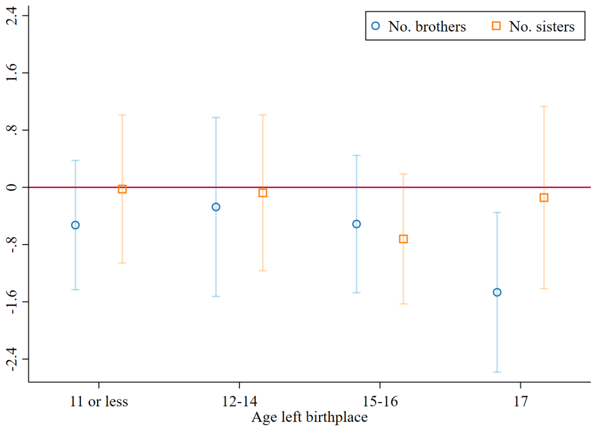
(a) Reason of migration



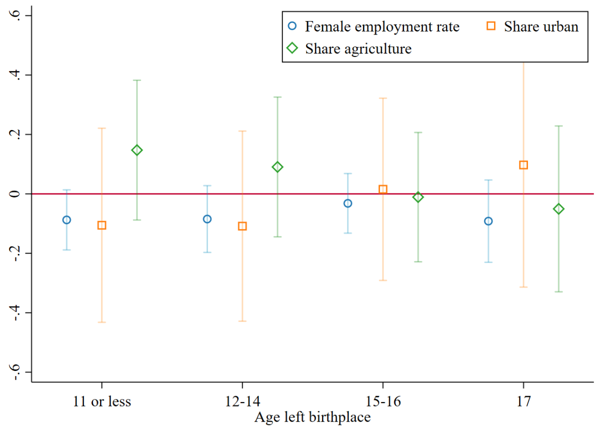
(b) Wealth proxies



(c) Number of siblings



(d) Characteristics of destination



Note: The figure show the coefficients on the interactions between age of emigration and birthplace female labor force participation. The regressions also control for (i) age of emigration dummies, (ii) birthplace female employment rate (iii) interactions between the emigration age and birthplace female employment rate. I chose 18 years old as the base category so that the interactions test whether the slopes at each age are different from that for women emigrating at 18. Data on reasons for emigrating is available only for people emigrating at 12 years old or older. Error clustered by regency of birth. The figure shows 90% confidence intervals. Data from the IFLS.

5 Robustness

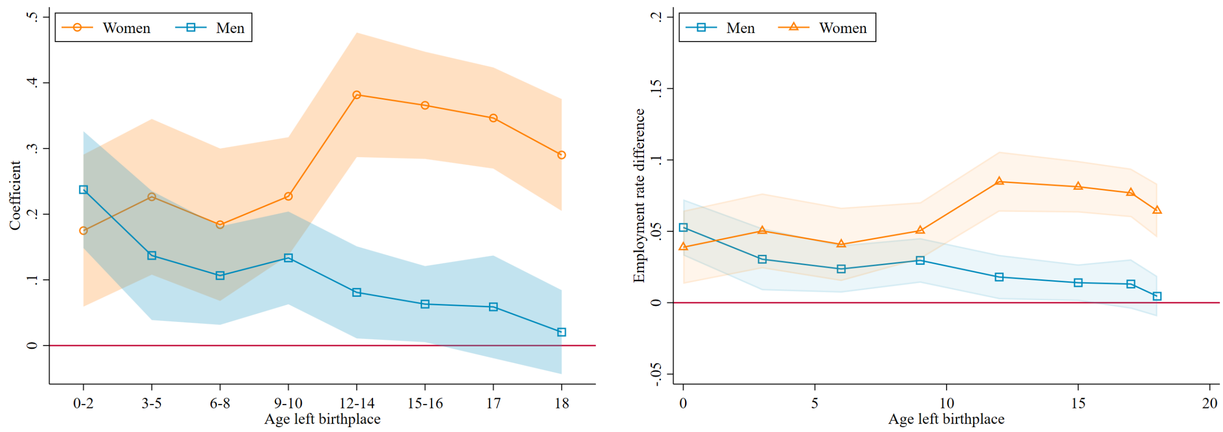
In this section I show that my results survive multiple robustness checks. I start by showing that I obtain similar conclusions when exploiting emigration age data from the Intercensal Survey. Next, I show my results are robust to (i) my choice of dependent variable, and (ii) the year from which I source the female employment rates.

5.1 I obtain similar results in alternative datasets

While very rich in information, the IFLS has two main limitations for the estimation strategy: (i) relatively small sample size, (ii) and lack of age information for emigres who left before they turn 12 years old. In figure 7 I take advantage of data from the 1985, 1995, and 2005 Indonesian Intercensal Surveys. This survey is similar to the census. It contains a more limited array of information than the IFLS. However, it has a larger sample size, and it contains information from which I can extract the exact age of emigration for all migrants. This additional information allows me to estimate place effects for earlier and smaller age brackets.

Figure 7: Indonesia: birthplace

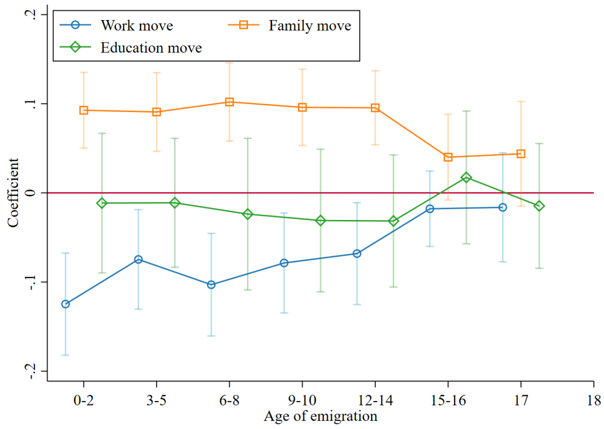
(a) Regression coefficients (b) Implied IQR gap



Note: Panel (a) shows estimates of the birthplace persistence coefficients by age of emigration $\boldsymbol{b_a}$. Panel (b) shows estimates of $\boldsymbol{b_a}\times$ regency-level IQR in female employment (22 p.p.). This is the implied gap in employment at different emigration ages between a person born in a regency at the 75th percentile in female employment, and another one born in a regency at the 25th percentile of female employment. The regression controls for a quadratic polynomial in age, education FE, and current regency FE. Shaded areas show 90% confidence intervals. The figure uses data from the 1985, 1995, 2005 Incercensal Surveys.

Panel A in figure 7 shows birthplace persistence estimates using the Intercensal Survey. Reassuringly, the results are similar in this alternative dataset. Women's birthplace persistence increases in the age of emigration and then levels off in their late teens. In contrast, men's persistence is mostly flat. Both patterns are consistent with a positive causal effect of birthplace that is exclusive to women.

Figure 8: Indonesia: estimates of birthplace slopes by age of emigration ($\boldsymbol{k_a}$)



Note: The figure shows estimates of the interactions between age of emigration dummies and the birthplace female employment rates. I chose 18 years old as the base category so that the interactions test whether the slopes at each age are different from that for women emigrating at 18. Sample restricted to women residing outside their birthplace. All regressions controls for year by bithplace regency fixed-effects, age of emigration fixed-effects, education fixed effects, and a quadratic polynomial on age. Standard errors clustered by regency of origin. Uses data from 1985, 1995, and 2005 Intercensal Surveys available in IPUMS international.

Nevertheless, there are two important differences relative to the IFLS results. Here, (i) selection is more important and thus I obtain smaller causal effects, and (ii) the causal effects level off slightly earlier. If we take the earliest point estimate figure 7 as the selection term, it is around 0.17. Although with a large standard error, this is about twice the size implied by the IFLS. The birthplace persistence goes up between 6 and 14 years old and then levels off at around 39 p.p. This implies smaller but still sizable birthplace effects. Panel (B) in the figure shows that the cumulative effect for two women on opposite sides of the IQR in female employment at 14 years old is of 4.6 p.p. This is 9% of the average female employment rate and it implies that 21% of the current spatial inequality is transmitted to the next generation of women.

The Intercensal survey also provides additional data to test the validity of the identification assumption. A potential concern with the tests in figure 6 is that these estimates are imprecise. Thus, I may not have the power to detect changes in selection with the IFLS sample sizes. To alleviate this concern, in figure 8 I show interaction estimates coming from the regressions where I use the emigration motive as dependent variable. This figure uses the larger samples from the Intercensal Survey. This figure suggests selection by emigration varies in late teens. Seventeen and eighteen year old women are more likely to be working than those at younger ages, so this change in selection is not surprising. However, the slope jumps at 15-16 years old are more concerning for the identification assumption. Note, however, that there is little evidence of changing selection between the ages of 6 and 14 years old. In fact, I cannot reject they are jointly equal. These are exactly the ages for which the Intercensal Survey estimates in figure 7 yield sizable place effects. Thus, in all, figure 8 supports the causal interpretation of the birthplace persistence estimates.

5.2 Results are similar for alternative measures of labor supply

All my main results use being employed as my main measure of labor supply. However, table 10 show that my main results carry through when I use alternative measures of labor supply. Column (1) shows my baseline estimates. In column (2) I use being a paid worker as an outcome. This adjustment is important because 23% of women are unpaid workers. It is thus reassuring that similar persistence patterns arise under this alternative measure. In columns (3) and (4) I use total weekly hours worked, and a dummy of being a full-time worker. I define full-time as working more than 35 hours per week. Because weekly hours information is not available in waves 4 and 5 of the IFLS, the samples for these two outcomes are substantially reduced. While admittedly less clear, the general persistence patterns are similar for these two outcomes: spending more time in high employment regencies increases women's labor supply in adulthood.

Table 10: Indonesia: birthplace persistence effects for alternative measures of labor supply

	Employed	**Paid worker**	**Weekly hours**	**Full time**
	(1)	(2)	(3)	(4)
Age of emigration interactions				
11- × p_b	0.09	-0.00	13.96	0.19
	(0.11)	(0.11)	(11.04)	(0.20)
12 × p_b	0.21	-0.01	15.03	0.28
	(0.15)	(0.15)	(10.94)	(0.21)
15 × p_b	0.53***	0.41***	10.84	0.14
	(0.11)	(0.11)	(8.93)	(0.17)
17 × p_b	0.57***	0.42***	20.84**	0.35*
	(0.13)	(0.14)	(8.44)	(0.19)
18-19 × p_b	0.54***	0.59***	31.97***	0.41**
	(0.13)	(0.12)	(9.34)	(0.18)
Observations	27,977	27,977	8,599	8,599
No. individuals	2,629	2,629	1,156	1,156
No. migrants	2,629	2,629	1,156	1,156
r2	0.16	0.13	0.17	0.15

Notes: I define full-time work as working more than 35 hours per week. Weekly hours data is not available for waves 4 and 5 of the IFLS. This substantially reduces the sample in columns (4) and (5). Table restricts the sample to people residing outside their birthplace with known age of outmigration. All regressions control for religion, and education FE; and a quadratic polynomial on age. Data from the ILFS.

6 Conclusions

In this paper, I document large and persistent spatial inequality in women's labor supply in Indonesia, a country with more than 118 million women. I argue that a substantial portion of this inequality is driven by the local environment women are born into. To identify the causal effect of place, I leveraged variation coming from the age women emigrated from their birthplace. I compared the labor supply choices of women who currently live in the same location, but who emigrated from their birthplace at different ages as children. If the omitted variable bias is independent of the age of emigration, this strategy allows me to distinguish the causal effect of place from variation driven by differences in women's unobserved characteristics.

I women's birthplace is particularly important during the formative childhood and teen years. Staying in a location at the 75th percentile of female employment between 6 and 16 years of age makes women between 4 to 10 percentage points more likely to work than those born in a location at the 25th percentile. These magnitudes mean that between 21 and 45 percent of the current spatial inequality in women's employment transmits to the next generation of women. Therefore, these women-specific place effects can be an important driver of the large and persistent differences in women's labor force participation within countries.

An important question left for subsequent work is what drives these place effects. The effects I uncover can be the consequence of continual exposure to gender norms less amenable to women's work (Boelmann et al., 2021). On the other hand, the expectation of low labor market prospects can lead women to underinvest in human capital accumulation. Future research should focus on distilling the mechanisms behind these large geographic disparities. The form of effective policies aiming to encourage women's inclusion on the labor market depends crucially on what the main mechanism behind these results is.

References

Alesina, A., Giuliano, P., and Nunn, N. (2013). On the origins of gender roles: Women and the plough. *Quarterly Journal of Economics*, 128(2):469–530. 4

Angelov, N., Johansson, P., and Lindahl, E. (2016). Parenthood and the gender gap in pay. *Journal of Labor Economics*, 34(3):545–579. 13

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics : an empiricist's companion.* Princeton University Press. 16

Autor, D. H. and Dorn, D. (2013). The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review*, 103(5):1553–1597. 2, 11

Badan Pusat Statistik (2019). *Survei Sosial Ekonomi Nasional (Susenas), 2012 Core, https://doi.org/10.7910/DVN/12TVW1.* Harvard Dataverse, V1. 5

Badan Pusat Statistik (2020). *Survei Sosial Ekonomi Nasional (Susenas), 2016 Kor, https://doi.org/10.7910/DVN/BOXZWU.* Harvard Dataverse, V2. 5

Blau, F. D. and Kahn, L. M. (2017). The gender wage gap: Extent, trends, & explanations. *Journal of Economic Literature*, 55(3):789–865. 7

Boelmann, B., Raute, A., and Schonberg, U. (2021). Wind of Change? Cultural Determinants of Maternal Labor Supply. *SSRN Electronic Journal*. 4, 5, 15, 32

Brinkhoff, T. (2022). Indonesia: Administrative Division (Provinces, Regencies and Cities), retrieved from https://www.citypopulation.de/en/indonesia/admin/. 12

Bryan, G. and Morten, M. (2019). The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia. *Journal of Political Economy*, 127(5):2229–2268. 4, 5, 8, 9

Charles, K. K., Guryan, J., and Pan, J. (2018). The Effects of Sexism on American Women: The Role of Norms vs. Discrimination. *SSRN Electronic Journal*. 2, 4, 15

Chetty, R. and Hendren, N. (2018a). The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects. *Quarterly Journal of Economics*, 133(3):1107–1162. 2, 3, 4, 15, 17, 38

Chetty, R. and Hendren, N. (2018b). The impacts of neighborhoods on intergenerational mobility II: County-level estimates. *Quarterly Journal of Economics*, 133(3):1163–1228. 2, 4

Chetty, R., Hendren, N., Jones, M. R., and Porter, S. R. (2020). Race and economic opportunity in the United States: An intergenerational perspective. 4

Chetty, R., Hendren, N., and Katz, L. F. (2016a). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, 106(4):855–902. 2

Chetty, R., Hendren, N., Lin, F., Majerovitz, J., and Scuderi, B. (2016b). Childhood environment and gender gaps in adulthood. In *American Economic Review*, volume 106, pages 282–288. 2

Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A., and Cutler, D. (2016c). The Association Between Income and Life Expectancy in the United States, 2001-2014. *JAMA*, 315(16):1750. 2

Compton, J. and Pollak, R. A. (2014). Family proximity, childcare, and women's labor force attachment. *Journal of Urban Economics*, 79:72–90. 15

Damm, A. P. and Dustmann, C. (2014). Does growing up in a high crime neighborhood affect youth criminal behavior? *American Economic Review*, 104(6):1806–1832. 2, 4

Farre, L. and Ortega, F. (2021). Family Ties, Geographic Mobility and the Gender Gap in Academic Aspirations. *SSRN Electronic Journal*. 15

Fernández, R. (2007). Alfred Marshall lecture women, work, and culture. In *Journal of the European Economic Association*, volume 5, pages 305–332. 16, 17

Fernández, R. and Fogli, A. (2006). Fertility: The role of culture and family experience. In *Journal of the European Economic Association*, volume 4, pages 552–561. 16

Fernandez, R. and Fogli, A. (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics*, 1(1):146–177. 4

Fogli, A. and Veldkamp, L. (2011). Nature or Nurture? Learning and the Geography of Female Labor Force Participation. *Econometrica*, 79(4):1103–1138. 4

Hamory, J., Kleemans, M., Li, N. Y., and Miguel, E. (2021). Reevaluating Agricultural Productivity Gaps with Longitudinal Microdata. *Journal of the European Economic Association*, 19(3):1522–1555. 5

Hicks, J. H., Kleemans, M., Li, N. Y., and Miguel, E. (2017). Reevaluating Agricultural Productivity Gaps with Longitudinal Microdata. Technical Report 9, National Bureau of Economic Research, Cambridge, MA. 8

Kleven, H., Landais, C., and Søgaard, J. E. (2019). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4):181–209. 13

Le Barbanchon, T., Rathelot, R., and Roulet, A. (2021). Gender Differences in Job Search: Trading off Commute against Wage. 15

Milsom, L. (2021). The Changing Spatial Inequality of Opportunity in West Africa. 4

Minnesota Population Center (2020). *Integrated Public Use Microdata Series, International: Version 7.3 [dataset].* MN:IPUMS. 5, 6, 7

Molina, T. and Usui, E. (2022). Female Labor Market Opportunities and Gender Gaps in Aspirations. *SSRN Electronic Journal.* 4, 15

Olivetti, C. and Petrongolo, B. (2014). Gender gaps across countries and skills: Demand, supply and the industry structure. *Review of Economic Dynamics*, 17(4):842–859. 7, 15

Olivetti, C. and Petrongolo, B. (2016). The Evolution of Gender Gaps in Industrialized Countries. *Annual Review of Economics*, 8(1):405–434. 7, 13

A The Empirical Strategy

When I introduce the age of emigration data in section 4.1, I assume that women's employment decisions are determined by place of residence fixed effects δ_c, age of emigration fixed effects λ_a, female labor force participation at birthplace p_b, and an error term η_i:

$$e_{it} \quad = \quad \delta_c + \lambda_a + \boldsymbol{\sigma}_a p_b + \eta_{it} \tag{9}$$

The error term embodies factors that are potentially important in determining women's decision to work, but which I do not observe. These factors could be correlated with the woman's birthplace employment rate. For example, generally I do not observe whether a woman's mother worked. This variable is naturally correlated with the birthplace female employment rate.

To simplify the discussion, I write this model in its matrix form as follows:

$$E \quad = \quad F\boldsymbol{\omega} + P\boldsymbol{\sigma} + \boldsymbol{\eta}$$

here F is a matrix containing place of residence and age of emigration indicators, P contains interaction between the age of emigration fixed effects and FLFP at birthplace, $\boldsymbol{\omega}$ stacks the location and age of emigration fixed-effects, $\boldsymbol{\sigma}$ is a vector containing the age of emigration effects $\boldsymbol{\sigma}_a$, and η is error term vector.

My main interest is estimating the birthplace effects vector $\boldsymbol{\sigma}$ consistently. For simplicity I can express the model in terms of the birthplace effects and the unobserved components by residualizing it from the age and residency fixed-effects. Let $\tilde{Z} = I - F(F'F)^{-1}F'$. Then,

$$\tilde{E} \quad = \quad \widetilde{P}\boldsymbol{\sigma} + \tilde{\boldsymbol{\eta}}$$

now let us consider the OLS estimate of the birthplace effects $\hat{\boldsymbol{\sigma}} = (\widetilde{P}'\widetilde{P})^{-1}\widetilde{P}'E$. Note that from the above expression it follows that,

$$\hat{\boldsymbol{\sigma}} = \boldsymbol{\sigma} + (\widetilde{P}'\widetilde{P})^{-1}\widetilde{P}'\tilde{\boldsymbol{\eta}}$$

Therefore:

$$\begin{aligned} \text{plim } (\hat{\sigma}) \quad &= \quad \boldsymbol{\sigma} + \text{plim } \left[(\widetilde{P}'\widetilde{P})^{-1}\widetilde{P}'\tilde{\boldsymbol{\mu}}\right] \\ &= \quad \boldsymbol{\sigma} + \boldsymbol{\gamma} \end{aligned} \tag{10}$$

now let us examine the meaning of expression (10) in detail. This expression implies that the OLS estimate for the birthplace persistence at age a is the sum of two terms: (i) the birthplace effect effect at that age $\boldsymbol{\sigma}_a$, and (ii) a term that captures the correlation between the residualized employment rate at the origin and the residual $\boldsymbol{\gamma}_a$:

$$\boldsymbol{b}_a = \boldsymbol{\sigma}_a + \boldsymbol{\gamma}_a \tag{11}$$

A.1 Identification

Assumption 1 (constant selection):

Selection on unobservables does not depend on the age of emigration, that is $\boldsymbol{\gamma}_a = \boldsymbol{k}$

Assumption 1 essentially requires that correlation between women's –unobservable– characteristics and women's origin is the same no matter the age at which they move. To see this, note that γ is driven two components:

$$\boldsymbol{\gamma} = \text{plim}\ \left[(\widetilde{P}'\widetilde{P})^{-1}\right] \text{plim}\ \left[\widetilde{P}'\tilde{\boldsymbol{\eta}}\right]$$

The term plim $\left[\widetilde{P}'\tilde{\boldsymbol{\eta}}\right]$ captures the correlation between women's birthplace and the unobserved characteristics. This is transparent by examining the general term of the vector $\widetilde{P}'\tilde{\boldsymbol{\eta}}$:

$$\sum_{i=1}^{N} \widetilde{p^a_b}\tilde{\eta}_i \tag{12}$$

where p^a_b denotes the –residualized– interaction between the birthplace female employment rates and age of emigration dummies. By law of the large numbers, this element converges to:

$$\mathbb{E}\left(\widetilde{p^a_b}\tilde{\eta}_i\right)$$

Two sufficient but not necessary conditions for the constant selection to be satisfied are:

$$\mathbb{E}\left(\widetilde{p^a_b}\tilde{\eta}_i\right) = \boldsymbol{c} \tag{13}$$

$$plim(\widetilde{P}'\widetilde{P})^{-1} = Q \tag{14}$$

where Q non-singular matrix with: (i) diagonal elements equal to each other, and (ii) off-diagonal elements equal to each other.

The first condition requires the correlation women's unobserved characteristics and birthplace female labor force participation to be the same for women migrating at different ages as children. For instance, this condition allows for the fact that in places where more women work they were more likely to have working mothers. A violation of (14) would occur if, for example, women with working mothers stayed longer in their birthplace.

Moreover, note that equation (13) only contains variables residualized from current location and age of emigration fixed effects. It is likely that women of different ages migrate due to different motives. For example, women migrating at 10 years old would be less likely to migrate for school than 12 year old women because secondary in Indonesia starts at 13. However, this is not necessarily violates condition (13). This is only a problem if, after conditioning on current location and age of emigration fixed-effects, women migrating for family reasons came from different origins than women migrating for education. While I cannot fully test for this condition, I can provide supporting evidence by correlating birthplace female employment rate with observed women's characteristics for different emigration age cohorts.

Condition (14) imposes restrictions on the correlations between birthplace female employment for women migrating at different ages as children. It will be generally satisfied if women migrating at different ages came from roughly the same origins.

A.2 From OLS to causal effects

The constant selection assumption allows ne to identify the causal effects of spending more time at the birthplace. Identification follows the same intuition as in Chetty and Hendren (2018a). Because $\boldsymbol{\sigma}_a$ captures the birthplace effect accumulated up to age a, the effect of spending age a is just the difference across consecutive ages:

$$\boldsymbol{\pi}_a = \boldsymbol{\sigma}_a - \boldsymbol{\sigma}_{a-1}$$

under the constant selection assumption, equation (11) shows that simple subtraction of the OLS estimates identify the causal effects of place:

$$\boldsymbol{b}_a - \boldsymbol{b}_{a-1} = \boldsymbol{\sigma}_a - \boldsymbol{\sigma}_{a-1} \tag{15}$$

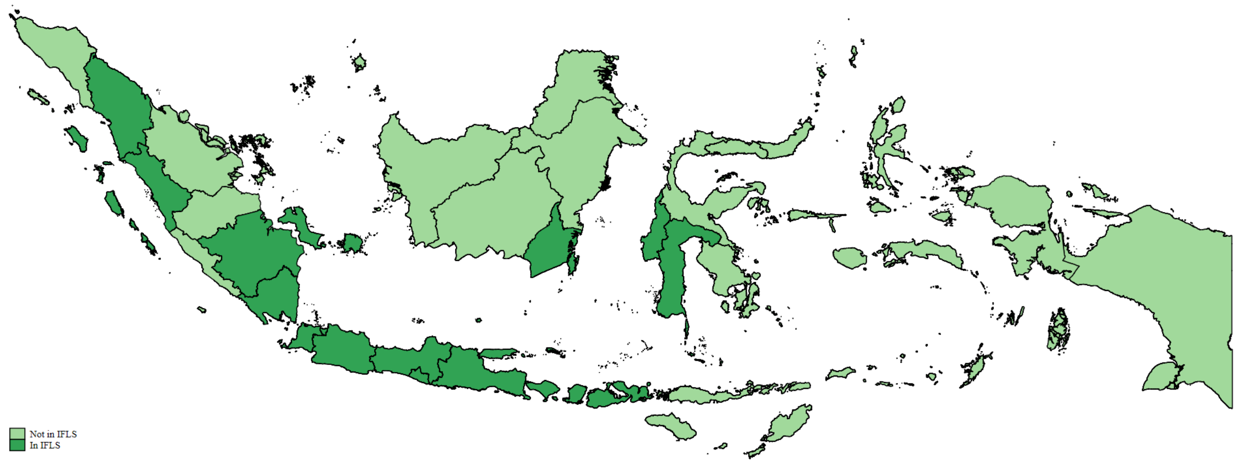
With an additional normalization, the OLS estimates can also identify the size of the selection term $\boldsymbol{\gamma}$. If we normalize the causal effect for the children with the least exposure to birthplace to zero ($\boldsymbol{\sigma}_0 = 0$), the OLS coefficient for this children is an estimate of the selection term:

$$\boldsymbol{b}_0 = \boldsymbol{c} \tag{16}$$

Equations (15) and (16) provide a full guide for estimating the causal effects. OLS estimates for women migrating at the earliest ages provide an estimate for the selection terms. Cross-age differences in the OLS estimates render the causal effect of spending a given age or period at the birthplace.

B Figures

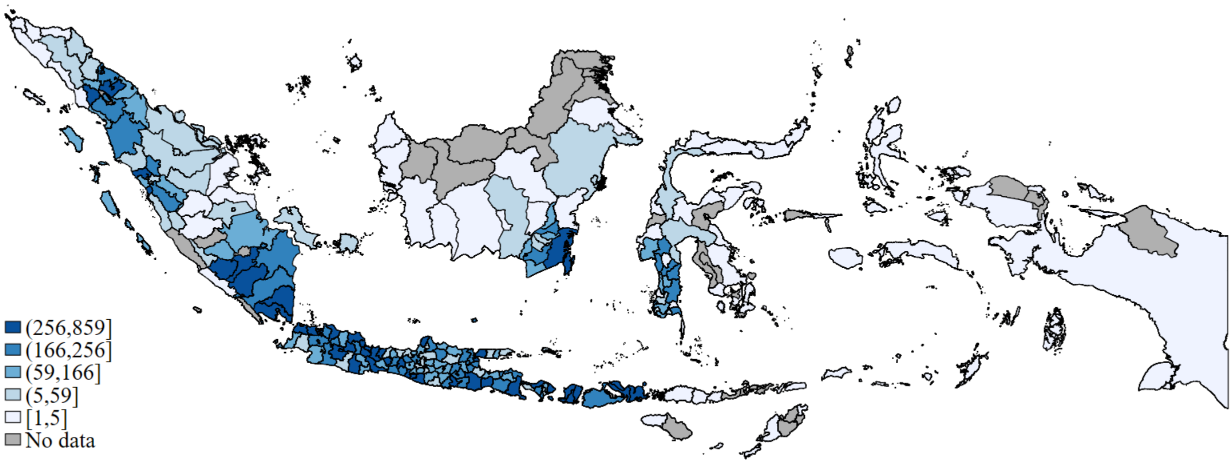
Figure B.1: Provinces in the original 1993 IFLS sample



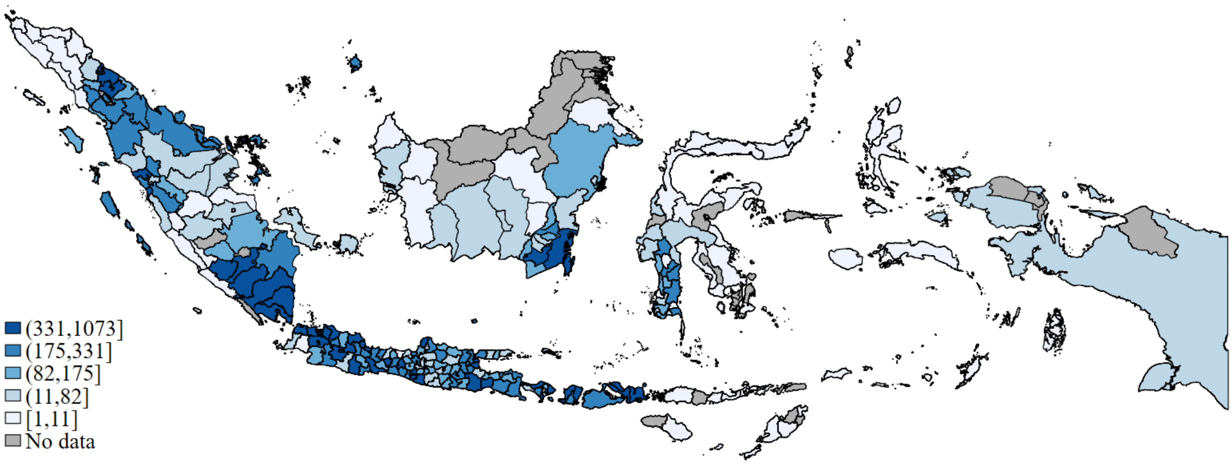
Note: The provinces from which the original 1993 IFLS sampled households. Because of migration, subsequent years can include individuals living outside these provinces. *Source:* RAND corporation.

Figure B.2: Indonesia: current and birthplace location of IFLS respondents

(a) Birthplace

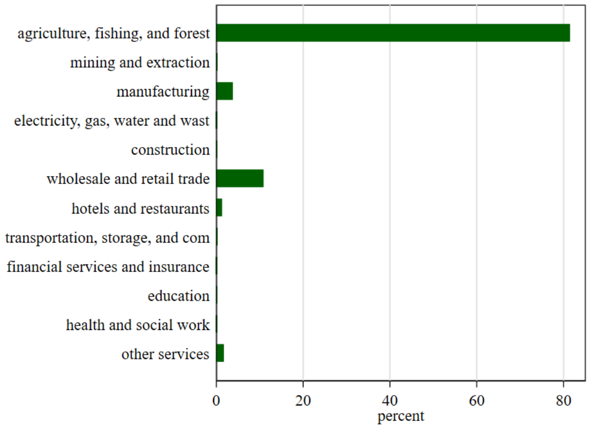


(b) Current location



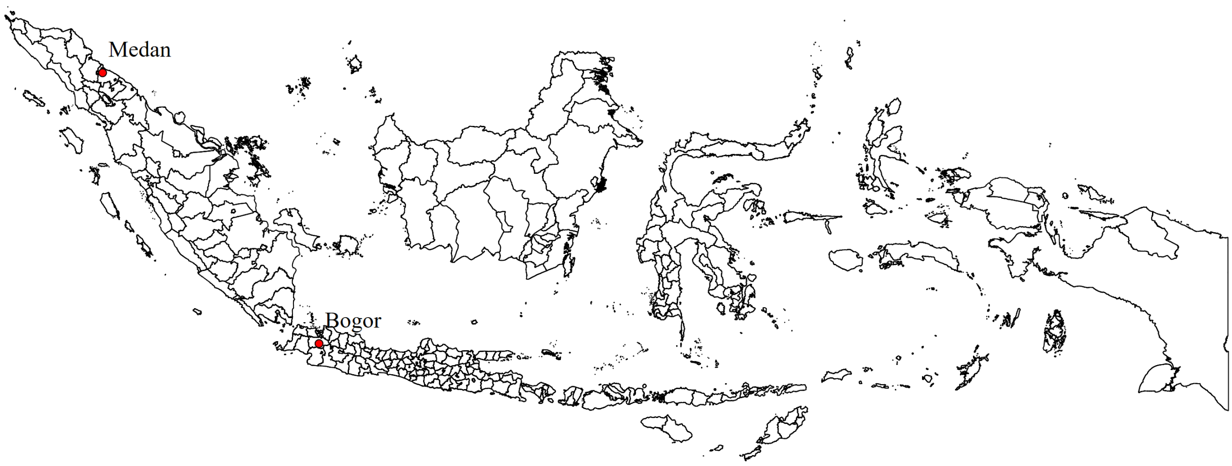
Note: The figure shows the counts of respondents by regency. It shows all the 268 regencies with consistent boundaries between 1970 and 2010. Each color groups a fifth of the regencies. The figure uses data from the IFLS.

Figure B.3: Indonesia: industry of employment of unpaid workers



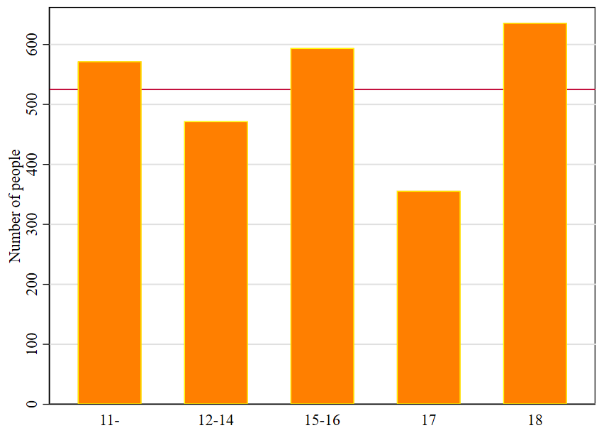
Note: Data from Indonesian Census 2010.

Figure B.4: Indonesia: location of selected regencies



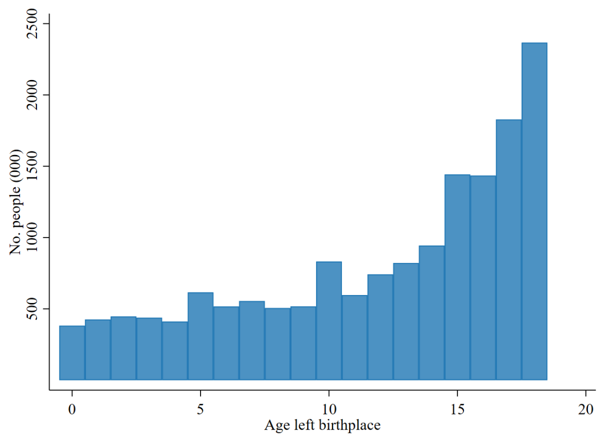
Note: The map shows with red dots the locations of the city of Medan and Bogor regency.

Figure B.5: Indonesia: number of women by age they left their birthplace



Note: Sample restricted to women who outmigrated before they turned 19 years old. The red line shows the average count across all age brackets. Data from the IFLS.

Figure B.6: Indonesia: number of emigres by emigration age, 0-19 years old



Note: Data from 1985, 1995, and 2005 Intercensal Surveys.

C Tables

Table 11: Indonesia: number of existing rengencies by year, 1980-2010

	1980	**1990**	**2000**	**2010**
Number of regencies	286	295	339	493

Notes: These regencies have chaning borders across decades. In my analysis, in each year, I aggregate these unis into 268 consistent-boundary regencies. Data IPUMS international.

Table 12: Regency-level summary statistics

	Mean (1)	**Std. Dev.** (2)	**Min** (3)	**Max** (4)	**Observations** (5)
Population	533,867	525,307	18,430	3,909,730	268
Share urban	0.45	0.30	0.07	1.00	268
Employment rate					
Women	0.53	0.14	0.29	0.91	268
Men	0.87	0.04	0.70	0.94	268
Industry composition					
Agriculture	0.43	0.23	0.00	0.81	268
Mining	0.01	0.03	0.00	0.26	268
Manufacturing	0.08	0.08	0.01	0.42	268
Construction	0.05	0.03	0.01	0.14	268
Services	0.35	0.13	0.12	0.68	268
Share married	0.75	0.05	0.56	0.93	268
Employment rate					
Share with at least high school					
Women	0.32	0.15	0.06	0.80	268
Men	0.36	0.15	0.10	0.82	268
Share literate					
Women	0.92	0.07	0.59	1.00	268
Men	0.95	0.05	0.72	1.00	268

Notes: table aggregates regencies to keep boundaries consistent across time. Table uses information from Indonesian Census and SUSENAS 2012.

Table 13: Autocorrelation in women's employment rate for different regency boundaries, 1980-2010

	Fixed boundaries			**Raw boundaries**		
	(1)	(2)	(3)	(4)	(5)	(6)
$t-10$ years	0.80			0.71		
	(0.02)			(0.02)		
$t-20$ years		0.72			0.56	
		(0.03)			(0.03)	
$t-30$ years			0.70			0.46
			(0.04)			(0.05)
Observations	800	534	268	878	541	278

Notes: sample restricted to women aged 18-64. Data IPUMS international. Robust standard errors in parenthesis. Columns (1) to (3) aggregate regencies to keep geographic units with consistent boundaries across time. Columns (4) to (6) the boundaries from the raw data.

Table 14: Indonesia: autocorrelation in regency-level women's employment rate, 1980-2010

	(1)	(2)	(3)
$t-10$ years	0.81		
	(0.06)		
$t-20$ years		0.73	
		(0.07)	
$t-30$ years			0.78
			(0.05)
Observations	800	534	268

Notes: The table shows the autocorrelation in regency-level employment rates. Data from IPUMS international. Robust standard errors in parenthesis.

Table 15: Indonesia: estimates birthplace persistence for alternative employment rate measures

Source of birthplace employment rates	(1)	(2)	(3)	(4)
Census 2010 (baseline)	0.36***			
	(0.04)			
Census 2000		0.29***		
		(0.04)		
Census 1990			0.27***	
			(0.04)	
Census 1980				0.26***
				(0.04)
Observations	64,727	64,727	64,741	64,727
N individuals	6,133	6,133	6,133	6,133
R^2	0.14	0.14	0.14	0.14

Notes: The table shows estimates of the birthplace persistence coefficients when I source the birthplace female employment rate from different census years. Uses data from IFLS and IPUMS international. Sample restricted to people residing outside their birthplace. All regressions controls for year, regency of residency, religion, and education FE, and a quadratic polynomial on age. Standard errors clustered by regency of origin.

Table 16: Indonesia: tets for no difference in selection in destination characteristics by age of emigration

	Childhood information					Characteristics of destination		
	Formal worker father	**11+ books**	**Ppl. per room**	**No. brothers**	**No. sisters**	**FLFP**	**% agriculture**	**% urban**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
p_b	-0.13	-0.26**	0.32	0.43	0.47	0.28***	0.10	-0.09
	(0.14)	(0.12)	(0.45)	(0.62)	(0.51)	(0.05)	(0.11)	(0.15)
Age of outmigration and p_b interactions								
11 or less	-0.08	0.17	0.01	-0.53	-0.02	-0.09	0.15	-0.11
	(0.21)	(0.18)	(0.41)	(0.55)	(0.63)	(0.06)	(0.14)	(0.20)
12-14	-0.30	0.23	-0.28	-0.27	-0.08	-0.08	0.09	-0.11
	(0.22)	(0.20)	(0.65)	(0.76)	(0.66)	(0.07)	(0.14)	(0.19)
15-16	-0.23	0.08	0.11	-0.51	-0.72	-0.03	-0.01	0.02
	(0.19)	(0.16)	(0.63)	(0.58)	(0.55)	(0.06)	(0.13)	(0.19)
17	-0.23	0.21	-0.07	-1.47**	-0.14	-0.09	-0.05	0.10
	(0.24)	(0.20)	(0.49)	(0.68)	(0.77)	(0.08)	(0.17)	(0.25)
Controls								
Emigration age FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE						✓	✓	✓
Age						✓	✓	✓
Age^2						✓	✓	✓
Interactions 0-17 are equal to zero								
F-statistic	0.62	0.49	0.15	1.82	0.69	0.68	0.54	0.40
p-value	0.65	0.74	0.96	0.13	0.60	0.61	0.71	0.81
Interactions 0-16 are equal to zero								
F-statistic	0.83	0.58	0.19	0.46	0.91	0.84	0.50	0.22
p-value	0.48	0.63	0.90	0.71	0.43	0.48	0.68	0.88
Observations	2,629	2,624	2,622	2,629	2,629	27,977	27,977	27,977
Number of people						2,629	2,629	2,629

Notes: This table shows test of the hypothesis that selection is constant across emigration age. It shows estimates of the interactions between age of emigration dummies and the birthplace female employment rates in regression (8). I chose 18 years old as the base category so that the interactions test whether the slopes at each age are different from that for women emigrating at 18. Columns (1) to (5) use one observation per individual. Standard errors clustered by birthplace regency. Data from the IFLS.

Table 17: Indonesia: tests for no difference in selection in reason of emigration

	Work (1)	Family (2)	Education (3)
p_b	0.11***	-0.08***	0.03
	(0.03)	(0.02)	(0.04)
Age of outmigration and p_b interactions			
0-2	-0.12***	0.09***	-0.01
	(0.03)	(0.03)	(0.05)
3-5	-0.07**	0.09***	-0.01
	(0.03)	(0.03)	(0.04)
6-8	-0.10***	0.10***	-0.02
	(0.03)	(0.03)	(0.05)
9-11	-0.08**	0.10***	-0.03
	(0.03)	(0.03)	(0.05)
12-14	-0.07*	0.10***	-0.03
	(0.03)	(0.03)	(0.04)
15-16	-0.02	0.04	0.02
	(0.03)	(0.03)	(0.05)
17	-0.02	0.04	-0.01
	(0.04)	(0.04)	(0.04)
Emigration age FE	✓	✓	✓
Destination regency by year FE	✓	✓	✓
Age	✓	✓	✓
Age^2	✓	✓	✓
Interactions 0-17 are equal to zero			
F-statistic	2.73	3.83	0.91
p-value	0.01	0.00	0.50
Interactions 0-17 are equal			
F-statistic	2.81	2.26	1.05
p-value	0.01	0.04	0.39
Interactions 3-14 are equal			
F-statistic	0.55	0.99	0.31
p-value	0.65	0.40	0.82
Observations	39,803	39,803	39,803

Notes: This table shows test of the hypothesis that selection is constant across emigration age. It shows estimates of the interactions between age of emigration dummies and the birthplace female employment rates in regression (8). I chose 18 years old as the base category so that the interactions test whether the slopes at each age are different from that for women emigrating at 18. Data from the 1985, 1995 and 2005 Intercensal Survey.