# Village Social structure and Labor Market Performance
## Evidence from the Philippines*

A. Stefano Caria†, Julien Labonne ‡

January 19, 2024

**Abstract**

This paper studies how social structure — the pattern of social links that connect individuals in a community — affects labor markets. Under competing views on the role of networks, social structures that discourage network hiring could improve or hinder labor market performance. We test these competing views using data on marriage networks in 15,000 villages, combined with labor force survey data. Using regressions analyses and an instrumental variable strategy, we find that individuals living in more socially fragmented villages are less likely to work in family firms, more likely to use formal job search strategies, invest more in education and earn higher wages. Social fragmentation thus discourages network hiring *and* improves labor market performance. These results survive 384 combinations of robustness checks. We further provide direct evidence against reverse causality.

**Keywords**: network structure, labor markets, job search

**Declarations of interest**: none.

*We thank Jonas Hjort, Clement Imbert, Jeremy Magruder and seminar and workshop participants at CAF, CSAE, JEES, LUMS, LSU, NEUDC, QMUL and SITE for feedback. All remaining errors are ours.

†University of Warwick. Email: `stefano.caria@warwick.ac.uk`, Web: www.stefanocaria.com.

‡University of Oxford. Email: `julien.labonne@bsg.ox.ac.uk`, `(Blavatnik School of Government University of Oxford Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG)`, Web: http://julienlabonne.wordpress.com.

## 1 Introduction

Social networks have a profound influence on the labor market (Granovetter, 2005; Beaman, 2016). Employers often hire using family networks and worker referrals (Heath, 2018; Chandrasekhar et al., 2020); jobseekers share information about job opportunities with their friends (Beaman and Magruder, 2012; Caria et al., 2020); migrants locate where their social connections are strongest (Munshi, 2020); and different types of peer effects can foster worker productivity and motivation (Mas and Moretti, 2009; Field et al., 2016). The way an individual's position in a social network determines their labor market outcomes has been studied both theoretically and in the field (Calvo-Armengol and Jackson, 2004; Bayer et al., 2008). On the other hand, the aggregate labor market implications of different network structures have rarely been documented empirically, and are generally not well understood.

There are two competing views on the role that network structure plays in the labor market. The first view emphasises the economic benefits that firms experience when they hire individuals from their social network. These benefits include better worker selection and stronger worker motivation (Montgomery, 1991; Heath, 2018). Under this view, network structures that make it easier for firms to hire from their social network should help firms increase productivity and, eventually, raise worker earnings. The second view, on the other hand, posits that network hiring, while rational from the point of view of the individual firm, is ultimately detrimental to labor market efficiency (e.g., Chandrasekhar et al. (2020)). If this view is correct, network structures that incentivize firms to hire from their social network lead to a worse allocation of talent and to weaker incentives for workers to accumulate human capital. This would in turn lower worker earnings — the opposite conclusion to the first view. Unfortunately, lack of data on social structure has so far prevented researchers from adjudicating between these competing views.

In this paper, we present novel evidence on how social network fragmentation — a key measure of network structure capturing the relative size of the clusters in which the network is divided — impacts labor market performance. We hypothesise that social fragmentation reduces employers' incentives to hire from their cluster of the network and provide empirical evidence supporting this hypothesis. We then turn to documenting the relationship between social fragmentation and labor market performance. Under the first view on the role of networks outlined above, social fragmentation should be associated with worse labor market outcomes. Under the second view, social fragmentation should be associated with better labor market outcomes. Our core contribution is thus to provide what is, to our knowledge, the first empirical test of these two views.

To carry out this test, we follow Cruz et al. (2020) and proxy social fragmentation with the likelihood that two randomly selected families in the village belong to different clans. We start by combining information on family names with local naming conventions to map out marriage networks in 15,000 villages in the Philippines. We then measure the social fragmentation of these marriage networks in two steps. First, in each village, we divide families into groups that have many connections with each other and limited connections to outsiders. For this purpose, we use some well-known community-detection algorithms (Girvan and Newman, 2002) to identify all clans/clusters in the village. The algorithm identifies groups of families with dense connections internally (i.e. within the group) and sparser connections between groups. Second, we compute the Herfindahl-Hirschman index of the clusters in each village, which gives us a measure of the likelihood that two randomly selected families in the village belong to different clans. In essence, this is a measure of the relative size of the clusters. In socially fragmented villages, average relative cluster size is small, and thus two randomly selected families are highly likely to belong to different clans.

We have two central empirical findings. First, we show that in more fragmented villages individuals are less likely to access the labor market through social networks. Specifically, a one-standard-deviation increase in social fragmentation is associated with a 15 percent reduction in the probability of working for a family-firm and a 10 percent increase in the probability of working for a private firm not owned by a relative. This evidence supports the hypothesis that fragmentation decreases incentives for network hiring.

Second, we document that social fragmentation is associated with higher hourly wages, longer working hours, and larger investments in human capital. These effects are stronger among women. Importantly, poverty levels are lower in more fragmented villages. A mediation analysis, using the method developed by Acharya et al. (2016), suggests that 64 percent of the boost in wages can be explained by changes in education while 25 percent can be explained by changes in the type of firm individuals work for. These results squarely support the second view on the relationship between network structure and labor market performance: social fragmentation discourages network hiring, and thus improves labor market performance through a combination of stronger human capital accumulation and a more efficient allocation of talent.

We implement a number of robustness checks that increase confidence in our findings. Importantly, we are able to control for a number of village-level characteristics that could be correlated with both fragmentation and economic outcomes, thus reducing concerns about omitted variable bias. Indeed, one may worry that fragmentation is correlated with urbanisation, population or other forms of social diversity. We show that our results are highly robust to including a rich set of controls capturing these potential confounders either on their own in the regression or all jointly. We also provide evidence against reverse causality — the hypothesis that economic outcomes change social fragmentation, e.g., by affecting migration or marriage patterns.[1] First, we show that social fragmentation is highly persistent and largely unrelated to a battery of municipal-level variables capturing education and assets in the 90s. To further reduce concerns related to migration, we include controls for how long individuals have lived in their villages, for the share of individuals who recently migrated into the village and the distance to the closest urban center. We also use an instrumental variable strategy that instruments current network links with old network links. Our results are unaffected. To reduce concerns related to changing marriage patterns, we show that an indicator of assortative matching in the marriage market has not changed over a 50-years time period. Finally, our specification curves show that our main results are robust to 384 combinations of our robustness checks.

Our results make three key contributions to the literature. First, to our knowledge, we are the first to provide evidence that can help adjudicate between two key competing views on the aggregate impacts of network structures on labor market performance. As mentioned above, despite an abundance of theory, empirical analyses of network structure are surprisingly rare. To circumvent the lack of field data on network structure, Dai et al. (2018) proxy network density with population density, Centola (2010) and Centola (2011) study online communities, and a number of researchers explore the effect of social structure in the lab (e.g. see Charness et al. (2014) and Gallo and Yan (2015)). Further, using the same data on family networks that we exploit in this paper, Cruz et al. (2020) investigate the effects of social structure on political outcomes.

Second, we show that social diversity can create economic dividends. Existing work has largely focused on the challenges posed by social diversity. For example, social rivalries can distort production (Hjort, 2014) and reduce support for redistribution (Alesina et al., 2018).[2] A smaller literature has documented an association between diversity and productivity, possibly due to skills complementarities (Ottaviano and Peri, 2006; Alesina et al., 2016). We advance this literature by proposing and providing evidence for a new channel through which an understudied dimension of diversity — social fragmentation — can affect economic performance.

Third, we highlight a novel mechanism that can distort the allocation of talent in

[1] Some of the interesting potential interactions between migration and social networks are explored in Munshi and Rosenzweig (2006) and Munshi and Rosenzweig (2016).

[2] Alesina and Ferrara (2005) provide an early review of this literature.

labor markets. The recent literature in development economics has devoted much attention to factor misallocation (Hsieh and Klenow, 2009). Many of the proposed explanations revolve around standard economic forces such as credit constraints and asymmetric information (Abebe et al., 2021b; Bandiera et al., 2017; Bassi and Nansamba, 2017; Dillon and Barrett, 2017; Abebe et al., 2021a). In this paper, on the other hand, we show that misallocation can also have social origins. This finding has important policy implications. In particular, it suggests that policy makers may be able to target interventions on the basis of the social structure of communities. In socially fragmented communities, removing credit constraints or providing information may be a viable and effective policy option. Similar policies are unlikely to have the same effects in more socially concentrated communities.

## 2 A simple conceptual framework

We present a simple conceptual framework that clarifies how social structure can impact labor market outcomes by changing the incentives to engage in network hiring, and thereby affecting worker effort and the allocation of worker talent. This framework shows how different mechanisms can generate competing predictions about the effects of social fragmentation on labor market performance.

We assume that in each village there is large number of workers; a fixed, small number of homogeneous firms, each managed by a single manager; and a set of bilateral social connections between individuals. As is common in real-life networks, social connections are organised in 'clusters' — groups of individuals who share many social connections with each other and have only limited connections with individuals outside the cluster. The village economy works in the following way. First, workers are randomly assigned a level of skills. Then, firms encounter random production opportunities over time. Once a production opportunity comes along, the firm manager hires an unemployed worker to produce output. 'Network hiring' occurs whenever the manager hires a worker who belongs to the same network cluster as the manager herself.

Firms have both a monetary and a social payoff. The social payoff is positive whenever the manager hires an individual who belongs to the same network cluster. If the manager hires outside of their network, the social payoff is zero. We assume larger clusters offer larger social payoffs, since larger clusters are often better placed to offer insurance, future opportunities, and social prestige.[3] The monetary payoff is determined

[3]This assumption is supported by several pieces of empirical evidence. For example, Munshi (2003) shows that individuals with larger social networks have better labor market outcomes; Angelucci et al. (2018) find that individuals redistribute government transfers through their networks and that larger

by worker skills and effort. The higher the skills of the worker, the higher the output they produce. Additionally, in large clusters where social pressure is high, a network hire will exert a high level of effort, which will raise output by a fixed amount. Network hires in small clusters and people hired outside of their network will exert low effort, which will not raise output. The worker is paid a wage that is given by the value of the output they produce minus some proportional rent for the firm. Hence firms maximise their absolute monetary payoff by hiring the worker that will produce the highest level of output.

In this framework, there are two reasons why network hiring may decrease as networks become more fragmented. First, the more fragmented the social structure, in the sense that average relative cluster size is small, the lower the average social payoff. This will induce managers to engage less frequently in network hiring. Second, the smaller the average cluster size, the lower is the chance that a network hire will exert high effort.[4] This will give a further incentive for firms to refrain from network hiring. Thus, we expect less network hiring in socially-fragmented villages.

Importantly, the reduction in network hiring impacts the allocation of talent and worker effort in opposite directions. As fragmentation increases and network hiring decreases, firms are more likely to hire the most skilled worker for the job, which improves the allocation of talent. At the same time, they become unable to motivate workers to exert extra effort through social network pressure. Depending on which channel dominates, firm output will increase or decrease as fragmentation rises. And changes in output will be reflected in worker earnings. This observation captures the core intuition behind our empirical test: a negative association between network fragmentation and worker wages provides support for first view outlined in the introduction — social structures that hinder network hiring limit labor market performance — while a positive association between network fragmentation and worker wages will provide support for second view — social structures that hinder network hiring boost labor market performance.

We can make an additional set of predictions by considering the possibility that individuals acquire new skills. In this case, network hiring will have a second cost: it will reduce incentives to invest in skills. This is because workers rationally anticipate that the employment gains that accrue from having stronger skills are lower due to

networks display a stronger overall consumption response to transfers; Ashraf et al. (2014) document that a prestige-based intervention has higher impacts among individuals who are part of a larger peer group.

[4]There is also a third effect at play. The higher the social fragmentation, the lower is the expected ability of the best candidate in the cluster.

network hiring. Lower skills will in turn depress output and wages. Thus, a positive association between social fragmentation and human capital investment will support the second view.

Finally, if the social payoff also differs by demographic characteristics, we expect to see that social fragmentation has stronger positive impacts for groups that offer lower social payoffs, and are thus more likely to face reduced opportunities due to network hiring. This could be the case for women, who typically have lower access to labor market networks. Our final prediction is that the impacts described above will be stronger for female workers.

## 3 Network measures and data

In this section we introduce the algorithm we use to measure social fragmentation and briefly present our various data sources. More information is available in the online appendix.

### 3.1 Measuring marriages

We follow Cruz et al. (2017) and use information on family names to measure family connections through marriage. This approach takes advantage of unique features of Filipino naming conventions: (i) within a municipality, a shared family name implies family connections; (ii) each individual carries two family names, which establishes that a marriage took place between members of those two families; (iii) names are difficult to change.

Names used in the Philippines were imposed by Spanish colonial officials in the mid-19th century. One of the stated objective was to distinguish families at the municipal-level to facilitate census-taking and tax collection. Last names were selected from the *Catalogo alfabetico de apellidos*, a list of Spanish names. They do not reflect pre-existing family ties. In each municipality a name was only given to one nuclear family. As a result, there is a lot of heterogeneity in names used at the local level, reducing concerns that names capture a similar ethnic background or other social grouping. Names are transmitted across generations according to well-established rules. Specifically, each individual has two family names: a last name and a middle name. A man's last name is his father's last name and his middle name is his mother's last name. Similar conventions apply to unmarried women. A married woman has her husband's last name and her middle name is her maiden name, *i.e.,* her father's last name.

We use data on individual's full names from the non-anonymized version of the National Household Targeting System (NHTS) (Fernandez, 2012) which was used to select beneficiaries for a large-scale conditional cash transfer programme. The names of all individuals in each village provides us with complete information on all marriages between families. We are thus able to reconstruct the full marriage network - with each family name being a node - in each village.[5] Importantly, as the department in charge of implementing the programme had to be able to reach out to selected beneficiaries, names were carefully inputted and cleaned. Importantly, it is often the case that multiple individuals in the village have the same combination of middle and last name (e.g, siblings, mother/children, etc.). Thus, even if there is a typo in the name of one individual, the relevant link between two families is still likely to be captured by the names of the siblings, children or mother of this individual, and hence it will be included in our map of the network.

### 3.2 Measuring social fragmentation

Our main empirical challenge is to measure social fragmentation at the village-level. Following Cruz et al. (2020), we measure how villages are divided into a number of clans. We rely on the notion of communities from social network analysis: groups of nodes with dense connections internally (i.e. within the group) and sparser connections between groups (Jackson, 2010). Those groups can serve as proxy for clans and we rely on the Girvan and Newman (2002) algorithm to identify them. The algorithm delivers a partition of C groups (indexed by $c = 1, \ldots, C$), each containing a share $s_c$ of nodes which represent families in our analysis. Our main measure of fragmentation is a standard Herfindahl-Hirschman index:

$$\texttt{fragmentation} = 1 - \sum_{c=1}^{C} s_c^2$$

where $s_c$ is the share of nodes in each clan $c$ and $C$ is the total number of clans. The measure captures the likelihood that two randomly selected families belong to different clans. As shown in Figure A.1, the levels of social fragmentation are high in our sample. To simplify interpretation we normalise the measure to be mean zero and standard deviation 1. To be clear, an increase in the measure captures an increase in the level of fragmentation. More details on how the communities are identified and the Herfindahl-Hirschman index is computed are available in Cruz et al. (2020) who use the same approach to link social fragmentation and political outcomes.

[5]Fafchamps and Labonne (2017), Cruz et al. (2017), Fafchamps and Labonne (2020) and Cruz et al. (2020) all use the same data and the same name matching algorithm to reconstruct the full marriage networks in those villages.

### 3.3 Labor force data

We also use Labour Force Survey (LFS) data collected by the Philippine Statistics Authority (PSA). We have access to all 26 quarterly surveys in the period July 2003 to October 2009.[6] We combine information included in the survey to compute hourly wage and weekly earnings. Respondents also provide information on the type of firm they work for (Family, private). In addition, for the sample of individuals who are looking for work, respondents provide information on how they look for jobs. We use this information to code whether they look for jobs through their networks or through more formal strategies.

Finally, we use data from the 2010 Census of Population and Housing to generate measure of ethnic and religious diversity as well as measures of the availability of key public goods at the village-level for our robustness checks.

## 4 Results

In this Section, we present our empirical findings. First, we show that social fragmentation is associated with a more widespread use of formal job search methods and a higher likelihood of working for a non-family firm. Second, we document that social fragmentation is also associated with larger investments in formal education, higher wages, and lower village-level poverty. We use mediation analysis to show that a large part of these impacts could be explained by the changes in education and family-firm employment associated with social fragmentation. Finally, we discuss the robustness of our results. Most of our regressions are at the individual-level but for some outcomes, such as poverty rates, we estimate village-level regressions.

We study the relationship between social fragmentation and individual outcomes of interest with models of the following form:

$$\texttt{outcome}_{ijm} = \alpha + \beta \cdot \texttt{fragmentation}_{jm} + \kappa \cdot X_{ijm} + \gamma \cdot W_{jm} + v_m + u_{im}. \qquad (1)$$

We use a series of outcomes variables capturing whether individuals are employed in a family firm or private firm, as well as individuals' job search strategy, education levels, wages and hours worked.

The unit of observation $i$ is an individual in village $j$ in municipality $m$. $\texttt{fragmentation}_{jm}$ is the Herfindahl-Hirschman index of social fragmentation discussed above, normalised to have mean zero and standard deviation one; $X_{ijm}$ is a vector of individuals controls

[6] More information on the survey design is available at: http://www.census.gov.ph/data/technotes/notelfs_new.htm visited on 26 March 2012. More details are available in Franklin and Labonne (2019).

(usually age and gender) and $v_m$ is a municipal fixed effect (as villages are nested within municipalities). At first, $W_{jm}$ is a vector of village controls, including information on population, urbanisation, as well as information on the average length of stay in the village. We then include a number of additional measures to reduce concerns about omitted variable bias: those include measures of migration, remoteness, wealth and ethnic/religious diversity. We discuss those carefully in Section 4.4. We cluster standard errors at the village-level as it is the level at which fragmentation is computed.

## 4.1 Job search, family firms, and human capital

Our first result is that individuals in more socially fragmented villages are less likely to access the labor market through social networks, and that they invest more in human capital. We show the headline specifications for these results in Table 1. We also report a more detailed analysis of the education variables in Table A.1 and look at the heterogeneity of the main effects by gender in Table 2. [7]

We document that individuals in socially-fragmented villages are less likely to access the labor market through social networks in a number of ways. First, we show that, among those individuals who hold a job, a one-standard-deviation increase in social fragmentation is associated with a 15 percent reduction in the probability of working for a family-firm (3 percentage-points) and a 10 percent increase in the probability of working for a private firm not owned by a relative (2 percentage-points). Second, among those individuals who are unemployed and are searching for work, we show that the likelihood of using formal job search methods (e.g., applying to formal vacancies by depositing a CV or registering to a public or private employment agency) is 4 percentage point higher when social fragmentation increases by one standard deviation, while the likelihood of approaching firms informally decreases by a similar amount. This evidence suggests that network hiring is less common in villages where social networks are more fragmented.

We also document sizable impacts on human capital accumulation, which hold both for high and low levels of education. In particular, a one-standard-deviation increase in social fragmentation is associated with a significant 5 percentage point increase in the probability of having some college (a 40% increase over the mean probability of having some college education in the sample). To put this in context, in our sample, individuals who live in urban areas are 14 percentage points more likely to have come college education compared to individuals who live in rural areas. Thus, the social-fragmentation

[7] The results presented in that table are estimated with versions of equation (1) where $\texttt{fragmentation}_{jm}$ is interacted with the gender dummy.

effect corresponds to more than one third of the urban premium for college attendance. Similarly, we find that a gain of one standard-deviation of social fragmentation is associated with a 9 percent reduction in the probability of having no education and a 12 percent reduction in the probability of having only primary education.

Importantly, both the human capital and the labor market effects are significantly stronger among women (Table 2). In particular, the effect on working in a private firm is twice as large, and the reduction in the likelihood of working in a family firm is four times as large. Labor-market networks are widely thought to be disadvantageous for women (Beaman et al., 2018), so a strong move away from network hiring is likely to be particularly beneficial for this group.

Taken together, this evidence corroborates the hypothesis that social structure changes the incentives to rely on social networks in the labor market. And, in particular, that the labor markets of villages where some large clusters dominate social life tend to rely more on social networks, while the labor markets of villages where social structure is more diverse and fragmented engage more robustly in formal, competitive hiring processes.

### 4.2 Wages, income and poverty

Our second result is that individuals in more socially-fragmented villages earn higher wages and have higher incomes. These findings support the second view on the role of networks – social fragmentation promotes better labor market performance. We illustrate these findings in a set of individual-level regressions in Table 1. Further, we report the results of village-level regressions that use income data from the NHTS dataset in Table 3.

The magnitudes of these effects are large. In the individual-level regressions, we find that a a one-standard-deviation increase in social fragmentation is associated with a 23 percent increase in the total wages individuals earn in a week. This effect is a combination of a 15 percent increase in hourly wages and an average gain of 1.6 hours worked per week (over a mean of 36 hours).[8]

The results on hourly wages are particularly important as hourly wages are typically thought to be a measure of worker productivity. In our setting, differential selection into employment complicates the interpretation of these regressions (e.g. there are more hours worked in more fragmented villages). However, we note that it is often assumed that marginal workers and marginal hours worked have lower latent wages.

[8]Not every employed survey respondent reports a wage. Hence, we run the wage regressions on a sample that is smaller than the sample of all individuals in employment.

Thus, selection is likely to moderate the positive effects on hourly wages that we document and, if this is true, our results on hourly wages would offer a lower bound of the true effects of social fragmentation on productivity.

We have two village-level measures of welfare: average per capita income and poverty rates. To capture impacts on the measures of welfare, we estimate village-level regressions similar to equation (1). We find that the wage effects identified at the individual-level translate into an increase in per-capita income of about 6 percent, or .18 of a standard deviation. Importantly, these gains are broadly distributed. Poverty rates decline significantly (by 3.7 percentage points, or 6 percent against a mean poverty rate of 61 percent).

### 4.3 Can the changes in education and formal labor market participation explain the wage effects?

The second view on the role of networks posits that social fragmentation foster a more competitive, formalised labor market. This, in turn, increases worker productivity by enabling a better allocation of talent and by incentivising investment in human capital. The results we have just presented support both of these hypotheses. However, if this framework holds some truth, we should also expect that the changes in wages we documented are commensurate to the estimated changes in education and labor market participation. In this subsection, we provide evidence on this point by presenting the results of a mediation analysis.

To perform our mediation analysis, we compute the average controlled direct effect (ACDE) of social fragmentation on wages, following the methods outlined in Acharya et al. (2016). The method improves upon the common strategy of controlling for post-treatment variables and allows researchers to estimate the share of the overall effect that can be explained by a given mediator. The ACDE is the impact of a treatment when a given mediator is not allowed to respond to the intervention.[9] In our context, this amounts to estimating the effect of social fragmentation on wages if either education or the type of firm were not allowed to change as social fragmentation increases. A comparison of the original effect to the ACDE thus reveals the importance of a given mediator. If the candidate mediator plays an important causal role, the ACDE will be small compared to the original effect — since a change in the value of the mediator is necessary to generate a large part of this effect. On the other hand, if the candidate mediator does not play an important causal role, the ACDE will be close in size to the

[9]Estimating this quantity requires the assumption of sequential unconfoundedness — we have to rule out the presence of unobservables correlated with both the mediator and the outcome of interest.

original effect. We present the results of our analysis in Figure A.2. In this analysis, we proxy education with a set of dummies capturing different levels of education attainment, and we proxy the type of firm with a set of dummies capturing whether an individual works in a family firm, in a private non-family firm or in a government firm.

We find that both human capital and firm type appear to mediate a large share of the effect of social structure. In particular, changes in education mediate about 64 percent of the main treatment effect and changes in the type of firm mediate about 24 percent of the effect of social fragmentation on wages. In other words, the patterns of mediation are consistent with the mechanisms proposed by the second view.[10]

### 4.4 Robustness

In this Section, we present an extensive series of tests to address concerns about reverse causality and endogeneity, and to establish the robustness of our findings to a large number of sample and measurement checks. We start by focusing on our headline regression on weekly wages (Table 4) but also show the results for all other outcomes variables (Appendix Figure A.3 - A.11.). Those results are estimated with equation (1).

A key concern is related to reverse causality. Villages with strong labor markets might attract more migrants, which can in turn increase social fragmentation. Alternatively, labour markets conditions can potentially change marriage patterns, thereby affecting social fragmentation. We deal with these concern in three ways. First, we show that social fragmentation is highly persistent and that current social fragmentation is largely unrelated to economic outcomes in the 1990s. To show this, in Table A.2 we regress our measure of current fragmentation on (i) social fragmentation obtained on the networks of individuals aged 45 or older — a proxy for past social fragmentation — and (ii) municipal-level measures of education and assets from the 1990 Census. We find that past social fragmentation is highly predictive of current social fragmentation: the simple bivariate regression has an $R^2$ of 0.62. Province fixed effects and a battery of 14 regressors capturing education and assets in the 1990s do not offer any meaningful additional predictive power: the $R^2$ of the full model is 0.64, only marginally larger than the $R^2$ of the model that only includes past social fragmentation. Furthermore, none of the covariates capturing education in 1990 is significantly correlated with current fragmentation and, while some assets measures are significantly correlated with current fragmentation, the point estimates are very small. For example, a one standard deviation increase in the share of individuals with a car is associated with a 0.017 standard

[10]We are unable to explore mediation with our measure of formal search as we only have this information for individuals who are currently looking for a job.

deviation units increase in fragmentation. Overall, these regressions provide empirical evidence against the hypothesis that the economic outcomes we consider in this paper have a meaningful impact on village social structure.

Second, to further reduce worries related to reverse causality driven by changes in marriage patterns, we investigate the stability of assortative matching on education — a key dimension of marriage patterns. Using the individual-level data, we compute the correlation between the spouses' education levels for different age groups (20-30, 30-40, 40-50, 50-60 and 70+), tracing marriage market patterns for a period of over 50 years. We find no evidence of any meaningful change in assortative matching over this time horizon: the correlations in the different cohorts are virtually identical, ranging between .68 and .70.

Third, to further reduce worries related to reverse causality driven by migration, we control for the average length of stay in the village (a measure of migration) in all regressions in the paper, and expand this in a number of ways in this section. We start by excluding areas that are classified as urban as they are the most likely migration destinations. Then, we use data from the Census to control for the share of recent migrants, specifically the share of individuals who lived in a different municipality 5 years ago. We are also worried that remoteness might be driving our results and control for the distance to the closest urban center. Finally, we follow Cruz et al. (2020) and construct networks based on individuals aged 45 or older. These networks would mostly reflect marriage decisions made decades ago and thus the social fractionalisation measures based on these networks are less likely to reflect reverse causality. We use those measures as an instrument for current social fragmentation.

We also check robustness over alternative dimensions. First, there exist alternative algorithms to identify network clusters and we want to rule out that our results are specific to the algorithm that we selected. We thus also implement the walktrap algorithm developed by Pons and Latapy (2006). Second, while we control for population, we are concerned about potential non-linearity in the relationship between population and wages, which could be captured by our measure of fragmentation. We thus create dummies for each decile of the population distribution and check that our results are robust to adding those as controls. Third, our measure of fragmentation might be correlated with inequality as well as with ethnic and religious diversity and so we control for those variables as well. Finally, we test that are our results are robust to controlling for public goods provision (Cruz et al., 2020) and to dropping a region, ARMM, where naming conventions may apply more loosely.

In Table 4 we show the robustness of the coefficient of our headline regression on weekly wages changes when we separately implement the robustness checks above. The coefficient on social fragmentation remains positive, large and significant in all specifications. In addition, we show that our results are robust to using 384 different combinations of the adjustments presented in Table 4.[11] The specification curves for the total wage, education, labor market participation, hourly wages, hours, poverty and income in appendix Figure A.3 - A.11

## 5 Conclusion

In this paper, we provide evidence to adjudicate between two competing views on the influence of social structure on labor markets. The first view posits that social structures that discourage network hiring are detrimental to labor market performance. The second view posits instead that those social same structures promote better labor market performance. We find that social fragmentation — a key dimension of social structure — is associated with a weaker use of network hiring. This, in turn, boosts hourly wages and reduces poverty. These results squarely support the second view on the role of networks. To the best of our knowledge, they are first evidence on the effects of social network structure on labor market outcomes using a large sample of villages.

[11] We run our regressions with all possible combinations of: 3 samples (full, excluding urban, excluding ARMM), estimation (IV and OLS), two measures of fragmentation and 5 set of controls variables (recent migration, distance to urban centers population, public goods and diversity).

## References


Abebe, G., A. S. Caria, and E. Ortiz-Ospina (2021a). The selection of talent: Experimental and structural evidence from ethiopia. *American Economic Review 111*(6), 1757–1806.

Abebe, G., S. Caria, and E. Ortiz-Ospina (2021b). The selection of talent. experimental and structural evidence from ethiopia. *American Economic Review 111*(6), 1757–1806.

Acharya, A., M. Blackwell, and M. Sen (2016). Explaining causal findings without bias: Detecting and assessing direct effects. *American Political Science Review 110*(3), 512–529.

Alesina, A. and E. L. Ferrara (2005). Ethnic diversity and economic performance. *Journal of Economic Literature 43*(3), 762–800.

Alesina, A., J. Harnoss, and H. Rapoport (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth 21*(2), 101–138.

Alesina, A., A. Miano, and S. Stantcheva (2018). Immigration and Redistribution. *NBER Working Paper No. 24733*.

Angelucci, M., G. De Giorgi, and I. Rasul (2018). Consumption and investment in resource pooling family networks. *The Economic Journal 128*(615), 2613–2651.

Ashraf, N., O. Bandiera, and B. K. Jack (2014). No margin, no mission? a field experiment on incentives for public service delivery. *Journal of public economics 120*, 1–17.

Bandiera, O., R. Burgess, N. Das, S. Gulesci, I. Rasul, and M. Sulaiman (2017). Labor Markets and Poverty in Village Economies. *The Quarterly Journal of Economics 132*(2), 811–870.

Bassi, V. and A. Nansamba (2017). Information Frictions in the Labor Market: Evidence from a Field Experiment in Uganda. *Working Paper*.

Bayer, P., S. L. Ross, and G. Topa (2008). Place of work and place of residence: Informal hiring networks and labor market outcomes. *Journal of political Economy 116*(6), 1150–1196.

Beaman, L. (2016). Social networks and the labor market. In *The Oxford Handbook of the economics of networks*.

Beaman, L., N. Keleher, and J. Magruder (2018). Do Job Networks Disadvantage Women? Evidence from a Recruitment Experiment in Malawi. *Journal of Labor Economics 36*(1), 121–157.

Beaman, L. and J. Magruder (2012, December). Who gets the job referral? evidence from a social networks experiment. *American Economic Review 102*(7), 3574–93.

Calvo-Armengol, A. and M. O. Jackson (2004). The effects of social networks on employment and inequality. *American economic review 94*(3), 426–454.

Caria, S., S. Franklin, and M. Witte (2020). Searching with friends.

Centola, D. (2010). The Spread of Behavior in an Online Social Network Experiment. *science 329*(5996), 1194–1197.

Centola, D. (2011). An Experimental Study of Homophily in the Adoption of Health Behavior. *Science 334*(6060), 1269–1272.

Chandrasekhar, A. G., M. Morten, and A. Peter (2020). Network-Based Hiring: Local Benefits; Global Costs. *NBER Working Paper 26806*.

Charness, G., F. Feri, M. A. Meléndez-Jiménez, and M. Sutter (2014). Experimental Games on Networks: Underpinnings of Behavior and Equilibrium Selection. *Econometrica 82*(5), 1615–1670.

Cruz, C., J. Labonne, and P. Querubin (2017). Politician family networks and electoral outcomes: Evidence from the philippines. *American Economic Review 107*(10), 3006–3037.

Cruz, C., J. Labonne, and P. Querubin (2020). Social structures, electoral competition and public goods provision. *American Political Science Review 114*(2), 486–501.

Dai, R., D. Mookherjee, K. Munshi, and X. Zhang (2018). Community Networks and the Growth of Private Enterprise in China. *Working Paper*.

Dillon, B. and C. B. Barrett (2017). An Updated View of African Factor Markets. *Working Paper*.

Fafchamps, M. and J. Labonne (2017). Do politicians' relatives get better jobs? evidence from municipal elections in the philippines. *Journal of Law, Economics & Organization 33*(2), 268–300.

Fafchamps, M. and J. Labonne (2020). Family networks and distributive politics. *Journal of the European Economic Association 18*(4), 1697–1725.

Fernandez, L. (2012). Design and implementation features of the national household targeting system in the philippines. *World Bank - Philippines Social Protection Note No 5*.

Field, E., S. Jayachandran, R. Pande, and N. Rigol (2016). Friendship at work: Can peer effects catalyze female entrepreneurship? *American Economic Journal: Economic Policy 8*(2), 125–53.

Franklin, S. and J. Labonne (2019). Economic Shocks and Labour Market Flexibility. *Journal of Human Resources 54*(1), 171–199.

Gallo, E. and C. Yan (2015). Efficiency and Equilibrium in Network Games: An Experiment. *Working Paper*.

Girvan, M. and M. E. J. Newman (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences 99*(12), 7821–7826.

Granovetter, M. (2005). The Impact of Social Structure on Economic Outcomes. *Journal of economic perspectives 19*(1), 33–50.

Heath, R. (2018). Why do firms hire using referrals? evidence from bangladeshi garment factories. *Journal of Political Economy 126*(4), 1691–1746.

Hjort, J. (2014). Ethnic Divisions and Production in Firms. *The Quarterly Journal of Economics 129*(4), 1899–1946.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly journal of economics 124*(4), 1403–1448.

Jackson, M. O. (2010). *Social and Economic Networks*. Princeton University Press. Princeton University Press.

Mas, A. and E. Moretti (2009). Peers at work. *American Economic Review 99*(1), 112–45.

Montgomery, J. D. (1991, December). Social networks and labor-market outcomes: Toward an economic analysis. *American Economic Review 81*(5), 1407–18.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the us labor market. *The Quarterly Journal of Economics 118*(2), 549–599.

Munshi, K. (2020). Social networks and migration. *Annual Review of Economics 12*(1), 503–524.

Munshi, K. and M. Rosenzweig (2006, September). Traditional institutions meet the modern world: Caste, gender, and schooling choice in a globalizing economy. *American Economic Review 96*(4), 1225–1252.

Munshi, K. and M. Rosenzweig (2016, January). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review 106*(1), 46–98.

Ottaviano, G. I. and G. Peri (2006). The Economic Value of Cultural Diversity: Evidence from US Cities. *Journal of Economic geography 6*(1), 9–44.

Pons, P. and M. Latapy (2006). Computing communities in large networks using random walks. *Journal of Graph Algorithms and Applications 10*(2), 191–218.

## Tables for inclusion in the paper

Table 1: Social Fragmentation, Network Hiring, Education and Wages

| Variable | Obs. | Mean | Point Estimate |
|---|---|---|---|
| Type of firm: Family firm | 581,781 | 0.20 | -0.03*** |
| | | (0.40) | (0.01) |
| Type of firm: Private firm | 581,781 | 0.21 | 0.02*** |
| | | (0.41) | (0.01) |
| Job Search: Direct Approach | 9,987 | 0.81 | -0.04*** |
| | | (0.39) | (0.01) |
| Job Search: Formal Approach | 9,987 | 0.17 | 0.04*** |
| | | (0.37) | (0.01) |
| Some College | 1,318,552 | 0.13 | 0.05*** |
| | | (0.34) | (0.01) |
| Report Wage | 581,781 | 0.27 | 0.05*** |
| | | (0.45) | (0.01) |
| Log(Weekly Wage) | 161,169 | 6.59 | 0.23*** |
| | | (0.95) | (0.02) |
| Log(Hourly Wage) | 161,169 | 5.05 | 0.15*** |
| | | (0.74) | (0.02) |
| Hours | 591,429 | 36.25 | 1.61*** |
| | | (19.30) | (0.31) |
| Hours (if wage) | 161,169 | 41.60 | 2.45*** |
| | | (16.65) | (0.32) |

Notes: Each row reports the point estimate on our fragmentation measure from a different individual-level regressions with municipal fixed-effects. The dependent variable is a dummy equal to 1 if the individual works in a family firm (row 1), works in a private firm (row 2), looks for job through his/her network (row 3), looks for job through formal channels (row 4), has some College education (row 5), report a wage (row 6), the log of weekly wages (row 7), the log of hourly wages (row 8) and hours worked in the past 7 days (rows 9 and 10). In row 10, we restrict the sample to individuals who report a wage. Regressions control for village-level average length of stay in the village, village population, the number of distinct families in the village, whether the village is classified as rural, as well as variables from the LFS surveys: age, gender, survey month and survey year. Standard errors, clustered by village, in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 2: Heterogeneity by Gender

| | Some College | Family | Private | Log Weekly Wage | Log Hourly Wage | Hours |
|---|---|---|---|---|---|---|
| Frag. | 0.047*** | -0.017*** | 0.030*** | 0.19*** | 0.12*** | 1.30*** |
| | (0.006) | (0.003) | (0.008) | (0.023) | (0.017) | (0.285) |
| Frag.*Female | 0.0064*** | -0.047*** | -0.023*** | 0.11*** | 0.092*** | 0.84*** |
| | (0.002) | (0.010) | (0.006) | (0.027) | (0.020) | (0.333) |
| Mean (male) | 0.12 | 0.14 | 0.26 | 6.60 | 5.07 | 36.7 |
| Std. Dev (male) | 0.32 | 0.35 | 0.44 | 0.89 | 0.65 | 17.6 |
| Mean (female) | 0.15 | 0.31 | 0.13 | 6.57 | 5.04 | 35.5 |
| Std. Dev (female) | 0.35 | 0.46 | 0.34 | 1.06 | 0.89 | 21.9 |
| Obs | 1318552 | 581781 | 581781 | 161169 | 164461 | 591429 |
| R2 | 0.092 | 0.22 | 0.15 | 0.19 | 0.22 | 0.11 |

Notes: Results from individual-level regressions with municipal fixed-effects. The dependent variable is a dummy equal to 1 if the individual has some College education (Column 1), works in a family firm (Column 2), works in a private firm (Column 3), the log of weekly wages (Column 4), the log of hourly wages (Column 5) and hours works in the past 7 days (Column 6). Regressions control for village-level average length of stay in the village, village population, the number of distinct families in the village, whether the village is classified as rural, as well as variables from the LFS surveys: age, gender, survey month and survey year. All variables are interacted with gender. Standard errors, clustered by village, in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 3: Social Fragmentation and Household Welfare

| | Poverty Rate | per capita income |
|---|---|---|
| Frag. | -0.037*** | 908.9*** |
| | (0.007) | (163.395) |
| Mean | 0.61 | 15359.1 |
| StdDev | 0.21 | 5127.5 |
| Obs | 15853 | 15853 |
| R2 | 0.62 | 0.63 |

Notes: Results from village-level regressions with municipal fixed-effects. The dependent variable is the village-level poverty rate (Column 1) and average (predicted) per capita income (Column 2). Regressions control for village-level average length of stay in the village, village population, the number of distinct families in the village, whether the village is classified as rural. Standard errors, clustered by municipality, in parentheses. * p < .10, ** p < .05, *** p < .01.

Table 4: Social Fragmentation and Log(Weekly Wage): Robustness Checks

| | Exc. Urban | Recent Migration | Distance to Urban Center | IV | Flexible Pop. | Alt. Alg. | Other Frag | Facilities | Exc. ARMM |
|---|---|---|---|---|---|---|---|---|---|
| Frag. | 0.20*** | 0.22*** | 0.18*** | 0.31*** | 0.23*** | | 0.21*** | 0.22*** | 0.24*** |
| | (0.029) | (0.025) | (0.026) | (0.046) | (0.027) | | (0.026) | (0.025) | (0.026) |
| Frag. (alt) | | | | | | 0.10*** | | | |
| | | | | | | (0.021) | | | |
| Observations | 107564 | 161169 | 134799 | 161169 | 161169 | 161169 | 161169 | 161156 | 153860 |
| $R^2$ | 0.17 | 0.18 | 0.17 | | 0.18 | 0.18 | 0.18 | 0.18 | 0.17 |

Notes: Results from individual-level regressions with municipal fixed-effects. The dependent variable is the log of weekly wages. In Column 1, all areas classified as urban are excluded from the sample. In Column 2, regressions control for the share of the population that moved to the municipality between 2005 and 2010. In Column 3, regressions control for the distance to the closest urban center. In Column 4, we instrument the fragmentation measure with the fragmentation obtained on the sample of individuals older than 45. In Column 5, regressions control for population flexibly (a different dummy for each population decile). In Column 6, the measure of fragmentation is computed using communities identified with the Latapy/Pons algorithm. In Column 7, regressions control for gini as well as ethnic and religious fragmentation. In Column 8, regressions control for the availability of key public goods in the village. In Column 9, all villages in ARMM are excluded from the sample. * $p < .10$, ** $p < .05$, *** $p < .01$.

# Appendix
# For Online Publication

Table A.1: Social Fragmentation and Schooling Investment

| | No Schooling | Some Primary | Primary Grad | Some HS | HS Grad. | Some College |
|---|---|---|---|---|---|---|
| Frag. | -0.0097*** | -0.041*** | -0.017*** | 0.0071** | 0.011*** | 0.050*** |
| | (0.004) | (0.006) | (0.003) | (0.003) | (0.003) | (0.006) |
| Mean | 0.11 | 0.34 | 0.14 | 0.16 | 0.11 | 0.13 |
| StdDev | 0.32 | 0.47 | 0.35 | 0.37 | 0.31 | 0.34 |
| Obs | 1318552 | 1318552 | 1318552 | 1318552 | 1318552 | 1318552 |
| R2 | 0.071 | 0.044 | 0.044 | 0.012 | 0.026 | 0.091 |

Notes: Results from individual-level regressions with municipal fixed-effects. The dependent variable is a dummy equal to one capture the highest level of schooling : none (Column 1), some primary school grades (Column 2), graduated from primary school (Column 3), some high school grades (Column 4), graduate from high school (Column 5) and some college (Column 6). Regressions control for village-level average length of stay in the village, village population, the number of distinct families in the village, whether the village is classified as rural, as well as variables from the LFS surveys: age, gender, survey month and survey year. Standard errors, clustered by village, in parentheses. * p < .10, ** p < .05, *** p < .01.

Table A.2: Predicting Fragmentation

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Fragmentation (over 45) | 0.74*** | 0.70*** | 0.69*** | 0.69*** |
| | (0.028) | (0.033) | (0.034) | (0.034) |
| No Schooling (Male) | | | 0.048 | 0.61 |
| | | | (1.007) | (1.046) |
| Some Primary (Male) | | | -0.027 | 0.68 |
| | | | (0.933) | (0.967) |
| Primary Grad (Male) | | | -0.43 | 0.26 |
| | | | (0.991) | (1.021) |
| Some High School (Male) | | | 0.31 | 1.02 |
| | | | (1.310) | (1.330) |
| High School Grad (Male) | | | 0.039 | 0.30 |
| | | | (2.904) | (2.849) |
| Some Primary (Female) | | | -1.10 | -1.07 |
| | | | (0.680) | (0.671) |
| No Schooling (Female) | | | -1.02 | -0.85 |
| | | | (0.735) | (0.723) |
| Primary Grad (Female) | | | 0.30 | 0.22 |
| | | | (0.818) | (0.785) |
| Some High School (Female) | | | -1.66* | -1.78* |
| | | | (0.988) | (0.982) |
| High School Grad (Female) | | | -0.31 | -0.76 |
| | | | (2.414) | (2.461) |
| Radio | | | | 0.17* |
| | | | | (0.098) |
| TV | | | | 0.23 |
| | | | | (0.174) |
| Vehicle | | | | 0.64* |
| | | | | (0.390) |
| Phone | | | | 0.70 |
| | | | | (1.728) |
| Province FE | No | Yes | Yes | Yes |
| Observations | 14720 | 14720 | 14720 | 14720 |
| $R^2$ | 0.62 | 0.63 | 0.64 | 0.64 |



Notes: Results from village-level regressions. The dependent variable is our main fragmentation measure Standard errors, clustered by municipality, in parentheses. * p < .10, ** p < .05, *** p < .01.

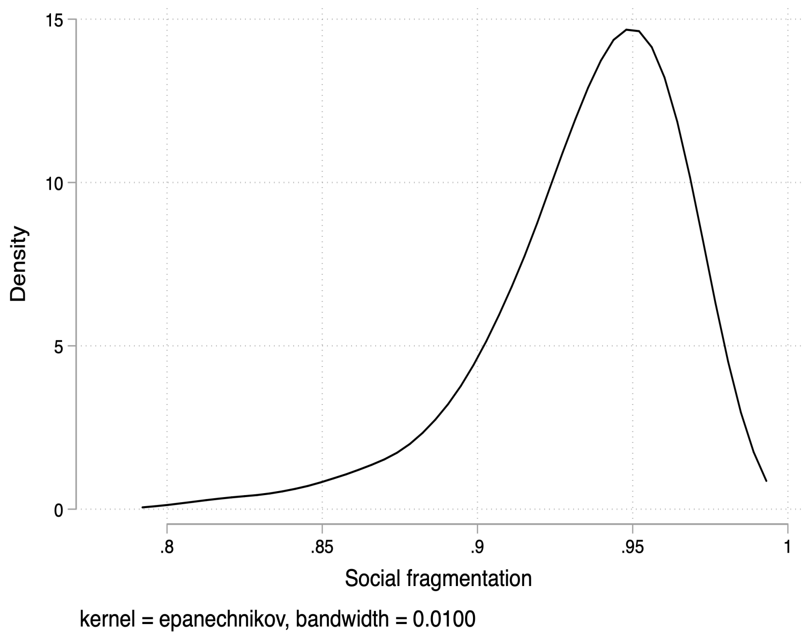


Figure A.1: Distribution of Social Fragmentation

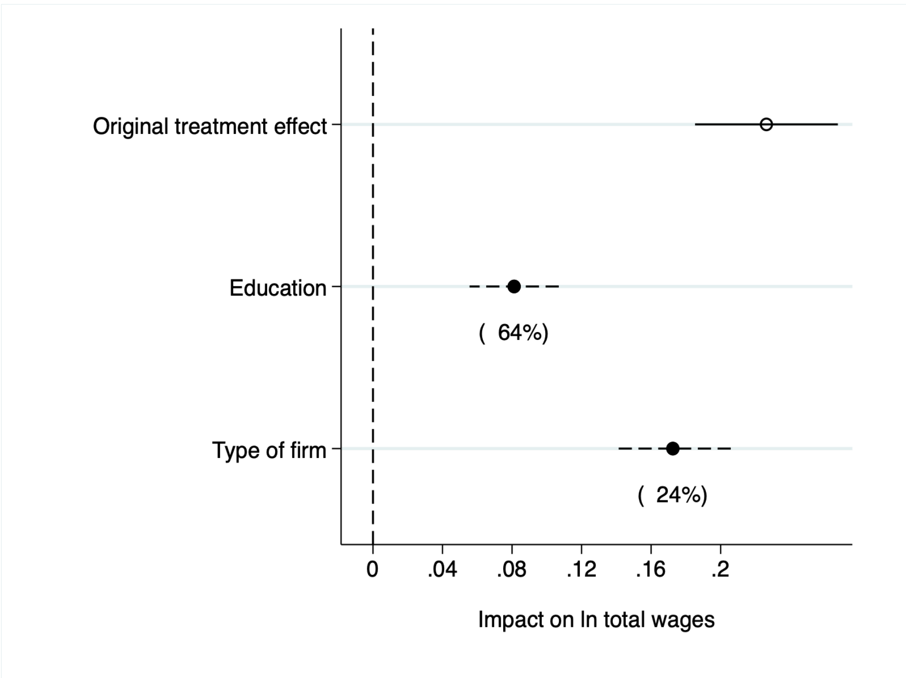


Figure A.2: Mediation Analysis

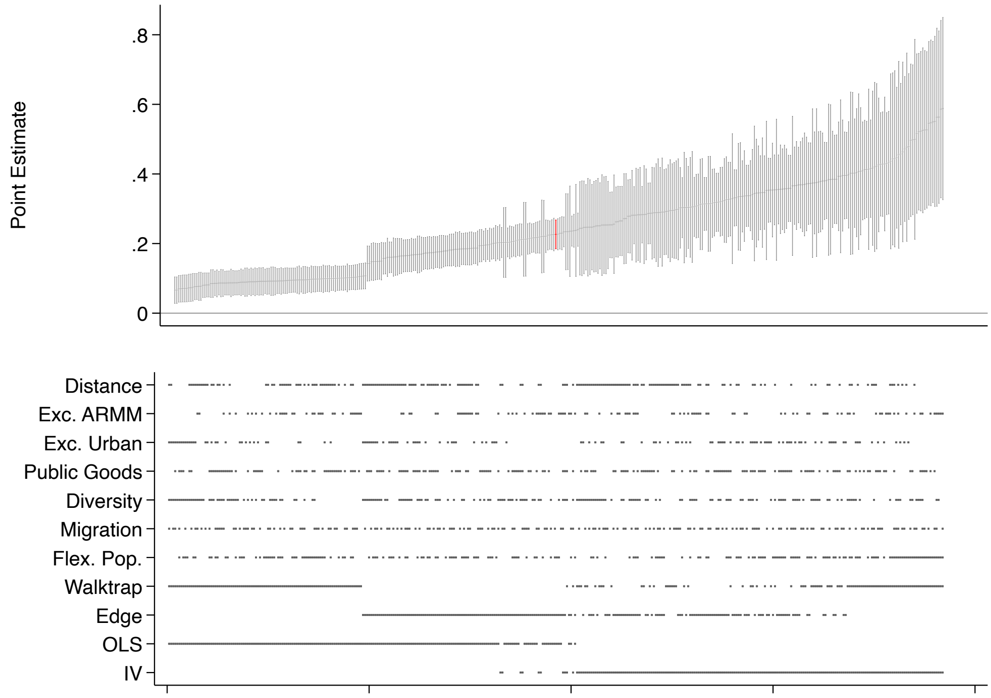


Figure A.3: Specification Curve for Log(Weekly Wage)

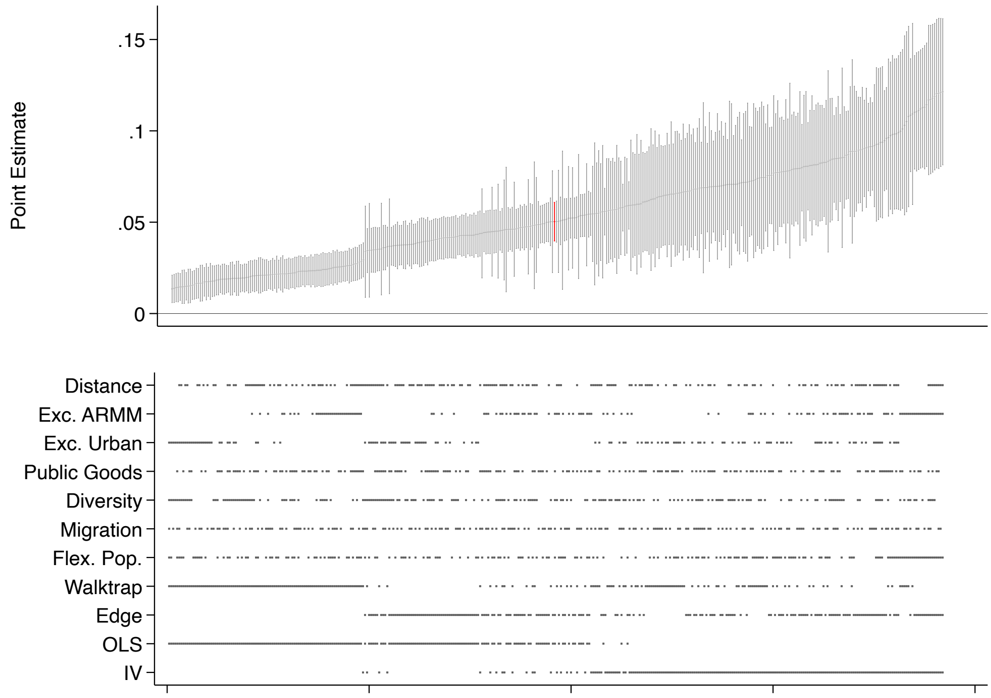


Figure A.4: Specification Curve for College

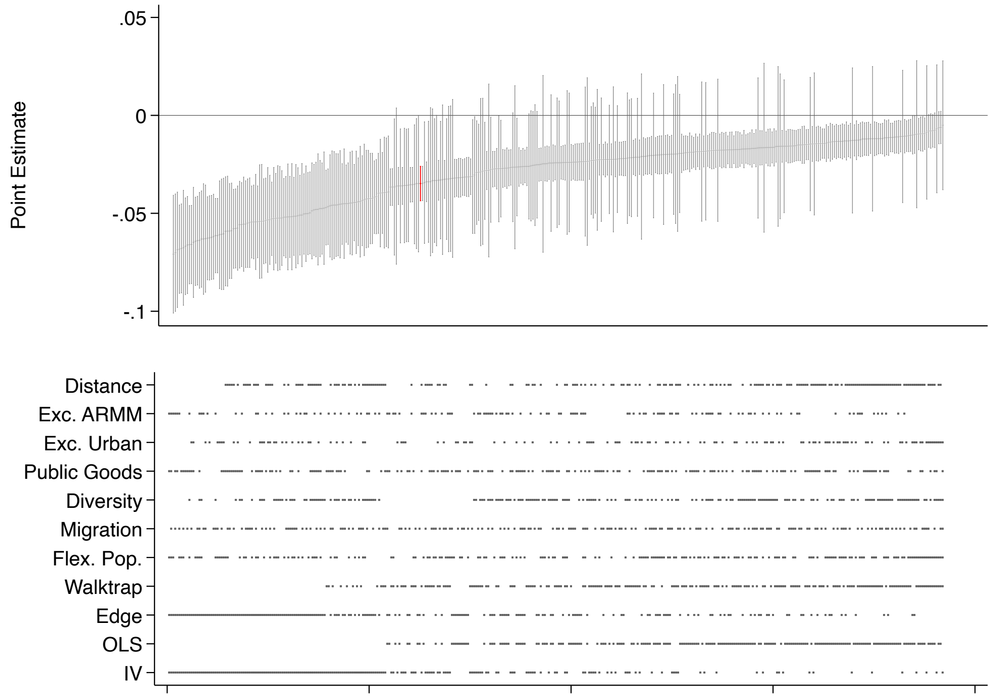


Figure A.5: Specification Curve for Family

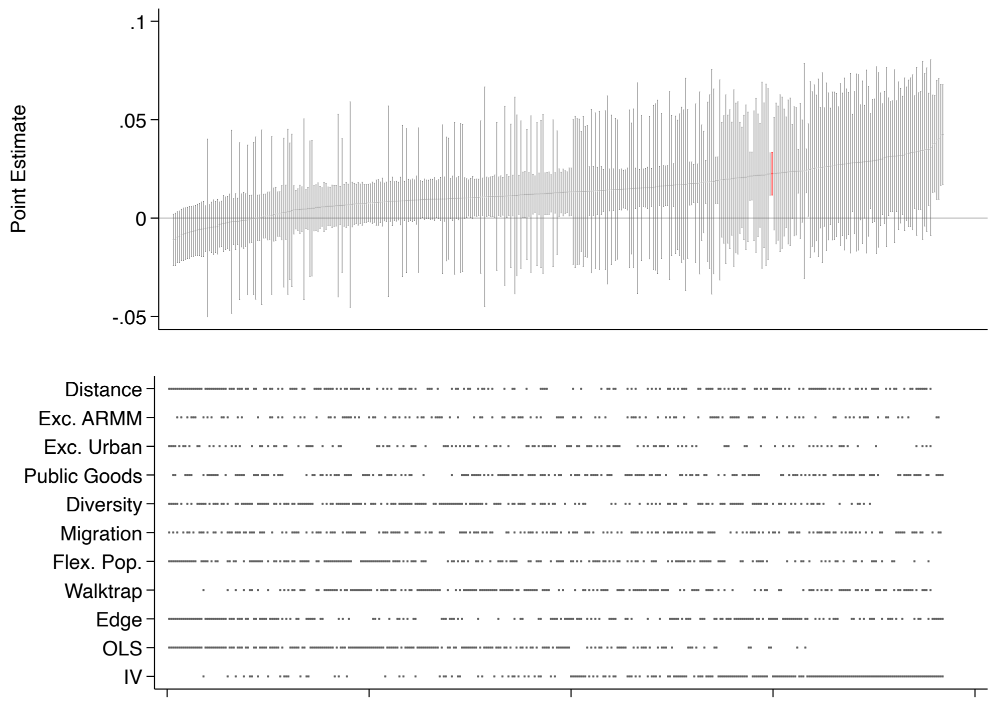


Figure A.6: Specification Curve for Private

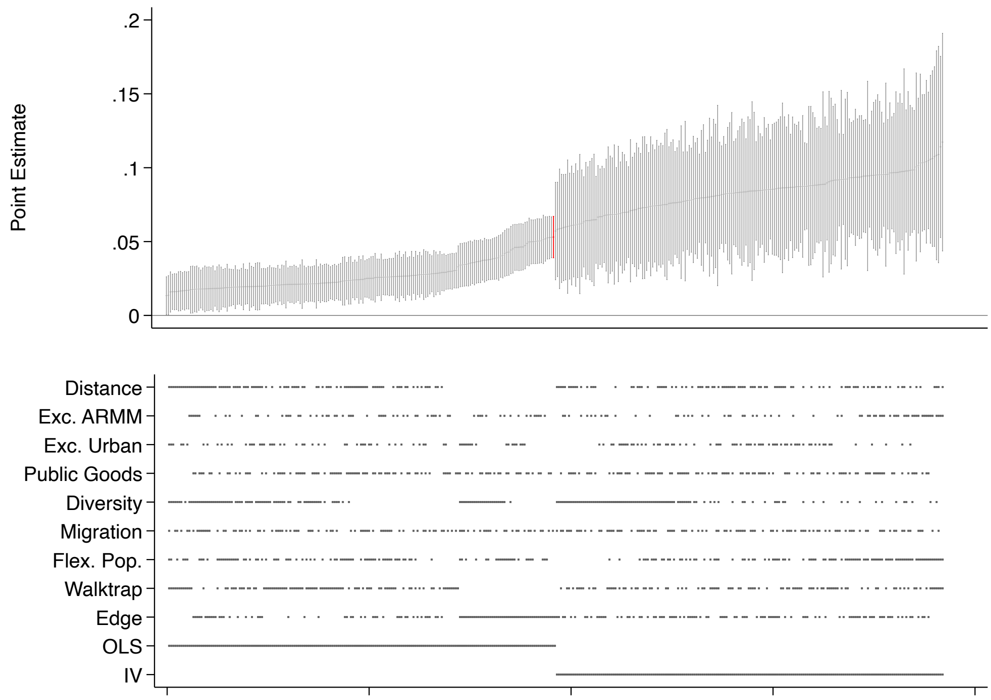


Figure A.7: Specification Curve for Report a Wage

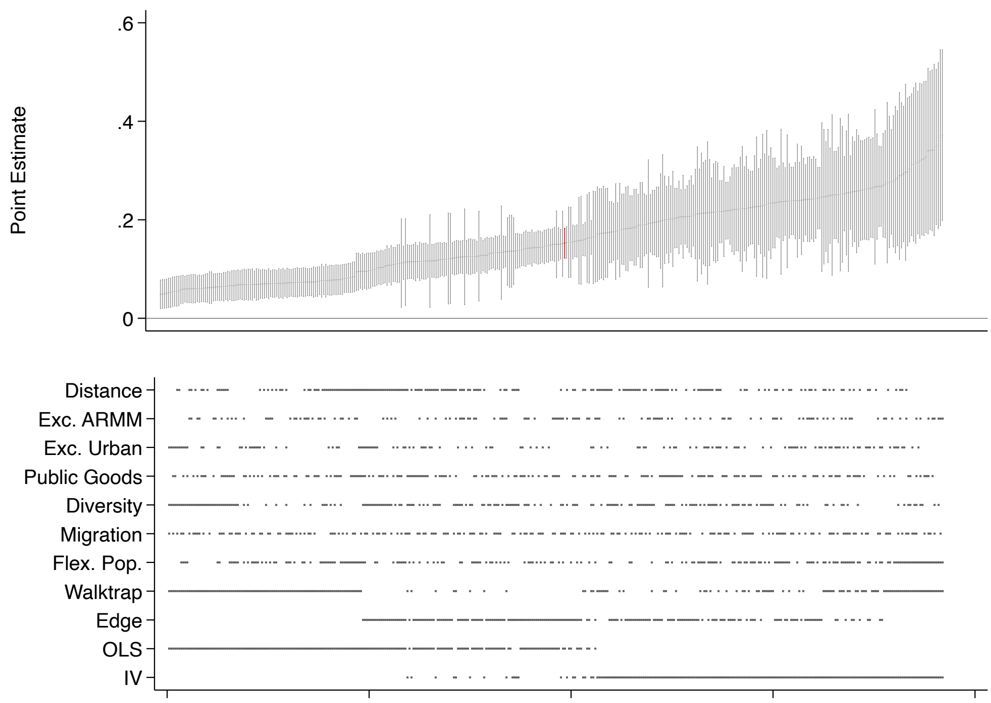


Figure A.8: Specification Curve for Log(Hourly Wage)

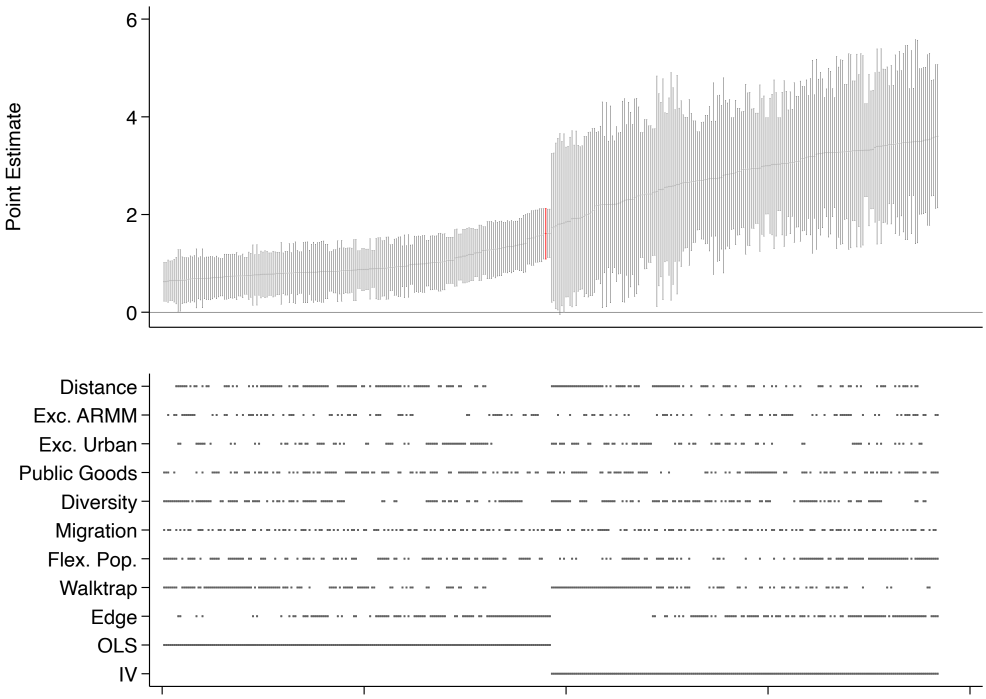


Figure A.9: Specification Curve for Hours

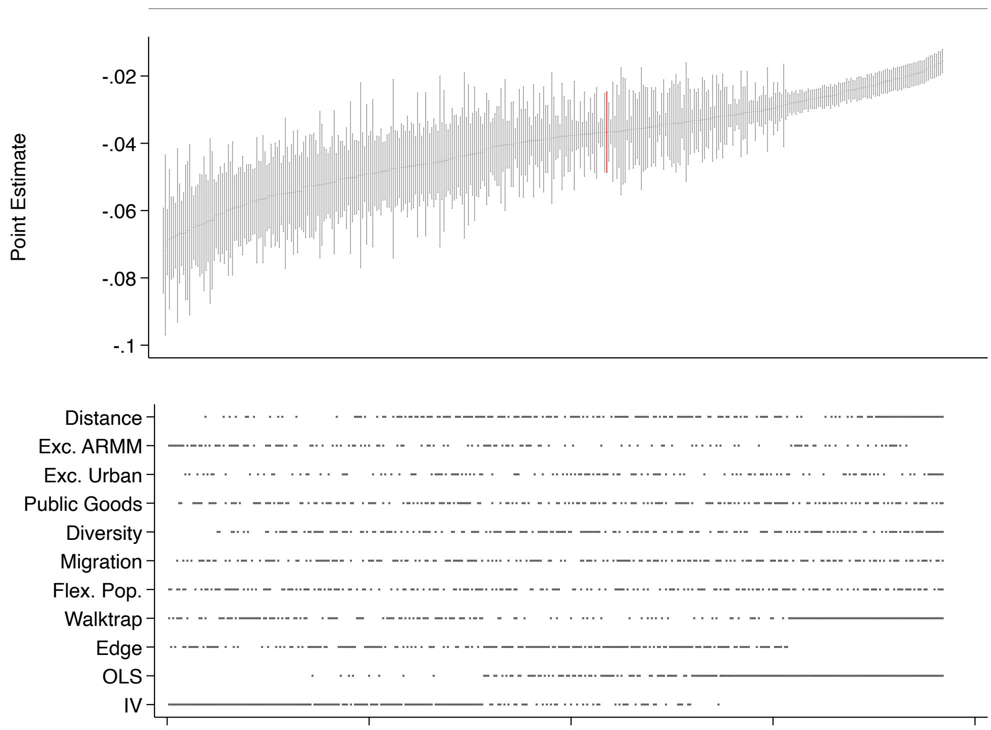


Figure A.10: Specification Curve for Poverty

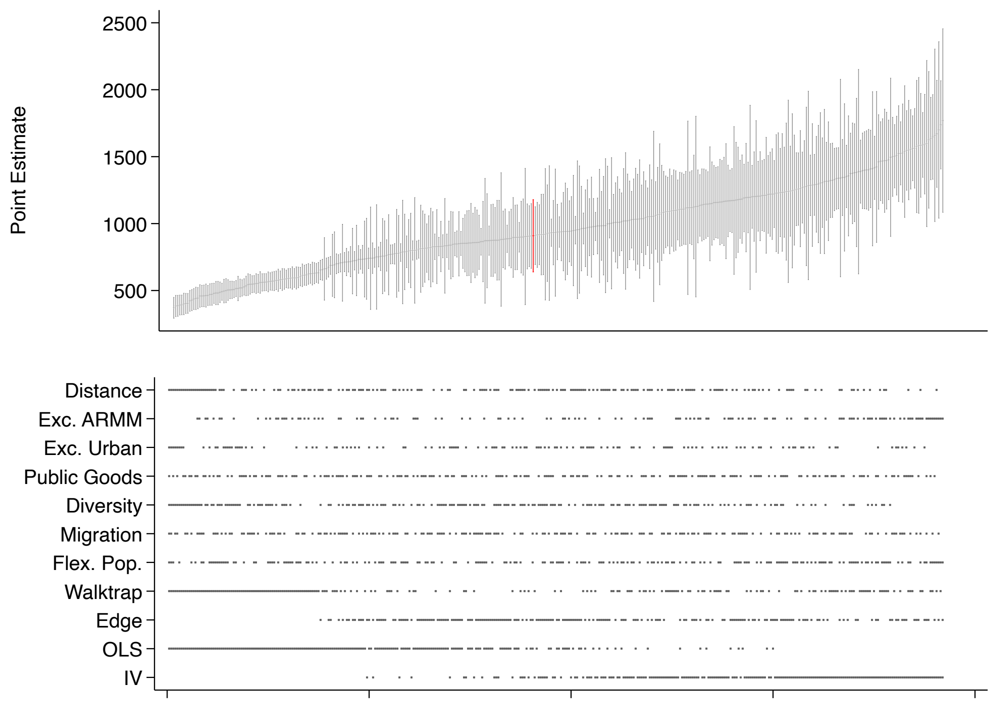


Figure A.11: Specification Curve for (predicted) per capita income