Durham University Business School

Department of Economics

Marriage Market Competition and Labor Supply

Joseph Pickens and Demid Getik

Working Paper No. 10, 2026

Durham Business School working papers series ISSN 1749-3641

Department Economics

Durham University Business School

Mill Hill Lane

Durham DH1 3LB, UK

Tel: +44 (0)191 3345200

https://www.durham.ac.uk/business/about/departments/economics/

Marriage Market Competition and Labor Supply*

Joseph Pickens
United States Naval Academy

Demid Getik
Durham University

July 13, 2026

Abstract

This paper examines how marriage market prospects influence decisions related to labor supply. Using U.S. microdata, we construct a novel and granular measure of marriage-market tightness. To address endogeneity from demographic mobility or labor-market shocks, we implement a Bartik-style instrument, approximating local population changes with national demographic shifts. We show that competition has heterogeneous effects across gender and education groups. For college-educated single men, greater competition increases hours worked and income. For non-college-educated single men, competition increases participation in the labor market, while the opposite holds for their female counterparts.

JEL codes: J12, J22, J21

Keywords: marriage markets; labor supply; labor force participation; American Community Survey (ACS); Bartik instrument

*We would like to thank Scott Kaplan, James Harrison, Aaron Sojourner, Patrizia Massner, Adrian Mehic, Alfonso Flores-Lagunes, Aleksei Chernulich, as well as audiences at the United States Naval Academy, LAC-PaL seminar series, Midwestern Economic Association Annual Meeting, Newcastle University, Sungkyunkwan University, University of Nebraska-Lincoln, the Swedish Institute for Social Research, and the Royal Economic Society Annual Conference, for their thoughtful comments and suggestions.

1 Introduction

Labor supply is one of the key indicators used by economists to assess an economy's state of well-being. The past several decades have seen a noticeable downward trend in labor supply among young men in the United States, and this trend has continued recently. According to data from the American Community Survey (ACS), the *non-participation rate* – the fraction of a population not in the labor force – for 18-24-year-old men has increased 34% from 2005 to 2019. While a significant portion of that change can be explained by increased uptake of tertiary education, that is likely not an exhaustive explanation. For 25-39-year-old men, who are less likely to be in school, non-participation rates increased 31% over the same period. The situation is not unique to the United States, with rates exhibiting similar trends in many other countries across the developed world.

A potentially important and understudied explanatory variable in this trend is the prospect of marriage and family formation. Supporting a family requires a significant investment of resources.[1] Therefore, not starting a family could remove a significant impetus for earning. Existing literature documents a substantial effect of diminishing work prospects on men's ability to form a family (Autor et al. 2019; Giuntella et al. 2022). Conversely, it is then plausible that changes in one's actual or perceived likelihood of forming a family could affect the motivation to earn and work. This is a particularly relevant concern given the sentiment and perception of the marriage market. For example, approximately two-thirds of men under the age of 30 in the United States have identified themselves as single, while the corresponding share for young women is closer to one-third (PRC 2023). Moreover, the share of never-married men between 25 and 39 increased by almost half between 2005 and 2019, according to the ACS. This increase in never-married men and the increased non-participation of those men can account for the entire 34% increase in rates mentioned above.[2] If family

[1] According to a 2015 estimate from the U.S. Department of Agriculture, raising a child from birth through age 17 costs $233,610 (in 2015 dollars, for a middle-income married-couple family with two children) (Lino et al. 2017).

[2] All else held equal, the increase in never-married men raises rates because they do not participate as frequently. Also, while the non-participation rate for never-married men increased by 29% over this period,

formation is an important determinant of labor supply and younger individuals perceive substantial changes in its likelihood, one could expect a noticeable effect on labor-related decisions.

In this paper, we examine whether changes in one's marriage prospects indeed have a direct effect on labor supply decisions. More specifically, we study whether increases in local marriage competition affect one's likelihood of non-participation in the labor force, hours worked, income, and likelihood of being *idle* – i.e., not participating in the labor market or education. Our analysis finds statistically significant results that vary across sex, age, and education. Despite less concerning participation trends, we examine young women alongside young men who also have skyrocketing never-married rates.[3]

The competition index – a novel measure of local marriage market tightness – uses national data on marriage patterns of 18-45-year-olds and local population data. While the framework can account for an arbitrary vector of characteristics, we focus on a *baseline measure* that accounts for marriage prospects by age; we also consider a measure with age and education in the appendix. The novelty of this measure is part of our contribution. Past work has mostly used the sex ratio to proxy for marriage market competition (e.g., Angrist 2002). Issues with the typical sex ratio include not removing those already married and the implicit assumption that all men compete with the same group of men for the same group of women (considering heterosexual men, for example). Our measure explicitly addresses these and other issues.

To address potential endogeneity, we instrument local demographic composition changes using the initial local population and national-level population changes. This Bartik-style approach should address concerns related to endogenous movements of individuals across different areas of the country, among other threats. The identifying assumption is that – conditional on age-location and location time fixed effects – initial demographic composition

it actually decreased slightly for those married before.

[3] In the 2005-2019 ACS data, female non-participation increased 6% among 18-24-year-olds and decreased 20% for 25-39-year-olds. Never-married rates for 25-39-year-olds increased 53%.

is unrelated to subsequent trends in local labor market outcomes except through its interaction with national demographic changes. Given the fixed effects, the identifying variation comes from differential changes over time across age groups within the same location.

We find that greater marriage market competition increases hours worked and earnings for college-educated men (age 25-39). For non-college educated men (of the same age), it decreases non-participation, and for younger men (age 18-24), it decreases idleness. A deviation from this trend is that hours worked decreases for those non-college men. Given the non-participation result, this may reflect a compositional change in the non-college labor force. As for non-college women, competition *increases* non-participation. Each of these estimates is generally robust to several methodological changes.

Our findings contribute to two main strands of literature. The first one broadly encompasses family formation and decisions pertinent to earnings and labor supply. Existing literature indicates the presence of a "marriage premium" for men, suggesting they earn more than their unmarried counterparts (de Linde Leonard and Stanley 2015; Ludwig and Brüderl 2018; Korenman and Neumark 1991). However, further evidence suggests that controlling for selection and productivity substantially reduces the size of the premium (McDonald 2020; Schechtl and Kapelle 2024). At the same time, increased economic opportunities appear to increase the number of births in affected men, even in the absence of an effect on marriages (Kearney and Wilson 2018). This literature also suggests a disparate child wage premium, with fathers earning more than childless men and mothers earning less than childless women (Kleven et al. 2019; Lundborg et al. 2017). However, more recent findings suggest that the "child penalty" tends to decline over time (Lundborg et al. 2024; Ali Akbari et al. 2025). We contribute to this literature by exploiting plausibly exogenous changes in the composition of the marriage market and observing the induced response in labor market decisions.

The second strand focuses more specifically on how economic conditions and earnings/wealth affect family formation. Cesarini et al. (2023) find that men who experience large positive wealth shocks through winning a lottery are more likely to get married and less likely to get divorced. Also, several studies indicate that external shocks to employment and earnings reduce marriage and fertility among males (Autor et al. 2019; Giuntella et al. 2022). Moreover, there is growing literature suggesting that a wife out-earning a husband can affect mental health in the couple (Getik 2024; Pierce et al. 2013; Johnston et al. 2023) and potentially increase the likelihood of divorce (Bertrand et al. 2015). It has also been associated with a higher quality of mates for women (Shenhav 2021), even if unmarried women seemingly have the opposite perception (Bursztyn et al. 2017). We contribute to this strand of literature by examining the opposite directionality of the effect, namely the impact of marriage prospects on one's labor supply and earnings. We therefore contribute to a better understanding of how labor market decisions might be affected by marriage and family prospects.

The paper is structured as follows. In Section 2, we overview the data and methodology used in the empirical analysis. In particular, we construct a general version of the marriage competition measure. Then, in Section 3, we present our results, and in Section 4, we offer a discussion. Finally, in Section 5, we conclude.

2 Methodology

This section details our empirical methodology. First, we overview our data source and explain the construction of our outcome variables. Then, we describe the construction of our marriage competition measure.

The measure is presented for an arbitrary set of characteristics, x, at the sex-location-year level. In our analysis, we focus on a single specification which only uses age. A supplementary specification considers age and education. We refer to these as the *baseline measure* and the *measure with education*, respectively. Both measures separate observations into single age bins from 18-45, and the latter measure uses a binary education measure where *college* indicates at least a bachelor's degree and *high school* indicates any other education level. We offer a brief justification of these two measures in Appendix A.1, but certainly, other specifications are possible and likely interesting. We leave such specifications for future work. Further, our analysis focuses on *single* adults in the United States – never married and not cohabitating with an unmarried partner.

Next, we give a general formulation of our OLS regression and explain why it might be biased. This potential bias motivates the use of an IV approach. To this end, we construct a Bartik-style instrument and present a 2SLS regression. Finally, we summarize a descriptive analysis of the competition measures given in Appendix A.3.

2.1 Data

The analysis relies on American Community Survey (ACS) microdata from 2005-2021 available through the Integrated Public Use Microdata Series (IPUMS 2025). The ACS is an annual survey conducted by the U.S. Census Bureau, collecting demographic information from 3-3.5 million households each year during our sample. We use individual-level data on respondents' sex, age, employment status, marital status and history, educational attainment, and geographic location, among other things.

One advantage of the ACS – relative to other publicly available surveys we could use, like the Current Population Survey (CPS) – is its granular geography. In particular, it provides the PUMA (Public Use Microdata Area) of each respondent. PUMAs divide each state into non-overlapping units containing at least 100,000 people. Since PUMAs change for each decennial Census, the ACS provides "consistent PUMAs" that are comparable across a larger number of years. This variable provides comparable values from 2005-2021.[4] Ultimately, this dictates our analysis period, but note that we excluded a base year (the first year, 2005) since the competition measure and its approximation are the same in that year. In the end, we use 1,078 geographical areas (i.e., consistent PUMAs). We refer to "consistent PUMAs" as "PUMAs" for convenience.

[4] In particular, we use the *CPUMA0010* variable. On the variable description page, they describe its construction as such: "we applied an aggregation algorithm that groups together 2010 PUMAs iteratively until the total population mismatch between each set of 2010 PUMAs and its closest matching set of 2000 PUMAs falls below 1% for both the 2000 and 2010 populations" (IPUMS 2025).

In the causal analysis, we focus on four outcome variables, each taken as an average in the corresponding cell.[5] To start, we remove those who do not report educational attainment or a labor force status. The first outcome is the non-participation rate. So for this outcome, $y^s_{a,\ell,t}$ is the fraction of sex-s, age-a respondents in location ℓ and year t who are out of the labor force. The second and third outcomes are hours worked (weekly) and income (annual), respectively. The fourth outcome is the rate of *idleness* – not being in school or working. ACS observation weights are used in the calculations of these cell-level statistics.

For these first three variables, we isolate those from age 25-39 who report not being in school at the time. For hours worked and income, we additionally remove those who earned less than $1,000 in 2017 U.S. dollars, and those who reported working less than 5 hours and more than 99 hours (who are top-coded) per week. For the fourth outcome, we isolate those aged 18-24 who report a school attendance status. Like others (e.g., Autor et al. 2019), we focus on this age range for idleness because it approximately covers the transition between school and work.

2.2 Construction of competition measure

Consider x to be a vector of individual characteristics. In the description that follows, we will use the baseline measure (i.e., $x = \{age\}$) as a running example to help build intuition.

Define $M_{x,\ell,t}$ to be the number of single men with characteristics x in location ℓ and year t. Let $W_{x,\ell,t}$ be the analogous number of single women. Next, define $\mathcal{P}^m_{x,x'}$ to be the probability of a single man with characteristics x – henceforth, an x-*man* – marrying a single woman with characteristics x' (i.e., a x'-*woman*) in the next year. It will be convenient to split this into the overall probability of an x-man getting married to a woman (of any characteristics) in the next year, γ^m_x, and the probability of them marrying a x'-woman conditional on them getting married, $P^m_{x,x'}$:

$$\mathcal{P}^m_{x,x'} \;=\; \gamma^m_x \;\cdot\; P^m_{x,x'}. \tag{1}$$

[5] Sex-age-location-year cell for the baseline measure and education-sex-age-location-year cell for the measure with education.

Figure 1: Conditional marriage probability distribution for 30-year-olds

Notes: For newlywed 30-year-old men, the red line gives the density of their spouses' ages. For example, around 4% of newlywed 30-year-old men marry a 32-year-old woman (we only consider heterosexual marriages here). The blue line gives the analogous density for 30-year-old women. While these are raw probabilities, the objects used in the analysis are "rescaled for availability" (see Appendix A.2 for details).

Let $\mathcal{P}^w_{x,x'}$, γ^w_x, and $P^w_{x,x'}$ be the analogous objects for women with $\mathcal{P}^w_{x,x'} = \gamma^w_x \cdot P^w_{x,x'}$.[6]

As an illustration of these distributions for the baseline measure, consider the distributions for 30-year-old men and women in Figure 1. While both distributions have a mode of 30, a 30-year-old newlywed man is more likely to have a younger wife, and a 30-year-old newlywed woman is more likely to have an older husband.

Now, consider a measure of competition that any single man faces for an x'-woman:

$$c^m_{x',\ell,t} \;=\; \sum_{x''} P^w_{x',x''} M_{x'',\ell,t} \Big/ W_{x',\ell,t} \tag{2}$$

[6]In the analysis, the objects $\mathcal{P}^m_{x,x'}$ and $\mathcal{P}^w_{x,x'}$ are rescaled for availability. The idea is that the raw probabilities are dependent on the age distributions of *available* (i.e., single) men and women. For example, in a counterfactual with more single older men, there would be more marriages of younger women to older men (ceteris paribus). See Appendix A.2 for technical details about the adjustment. Separately, it would be ideal to make these objects location-specific and perhaps time-specific as well. However, in the specifications of x we consider, we do not due to data sparsity.

First, note that this measure will be "close to 1" in the following sense. Using $x = \{age\}$ as an example, consider the measure for an age-a' woman (i.e., $x' = a'$). If populations are uniform across age and sex ($M_{a_1,\ell,t} = W_{a_2,\ell,t}$ for all ages a_1 and a_2) and the relative marriage probabilities are uniformly distributed ($P^w_{a',a''} = 1/(\#\text{ of ages})$ for every a''), then $c^m_{a',\ell,t} = 1$. Of course, these conditions are not true in general, and the measure and its average will vary.

This age-based measure has the flavor of a sex ratio but accounts for the fact that (heterosexual) marriage pairings of certain ages are more likely than others. For example, say we are talking about competition for women of age $a' = 35$. A relatively large population of 25-year-old men does not increase competition much since marriages between 35-year-old women and 25-year-old men are not very common. However, a relatively large population of 35-year-old men would (since marriages between 35-year-old women and 35-year-old men are more common).

With this "x'-specific" competition measure ($c^m_{x',\ell,t}$) in hand and its intuition, we now build up to the overall competition measure faced by an x-man, $C^m_{x,\ell,t}$. For simpler interpretation, it will be useful to define population shares. Let $m_{x,\ell,t}$ be the share of the population in location ℓ at time t that are x-men: $m_{x,\ell,t} = M_{x,\ell,t}/pop_{\ell,t}$, where $pop_{\ell,t} = \sum_{x'}(M_{x',\ell,t}+W_{x',\ell,t})$. Note that $pop_{\ell,t}$ only includes the groups we are considering, so $\sum_{x'}(m_{x',\ell,t} + w_{x',\ell,t}) = 1$ by construction. Let $w_{x,\ell,t}$ be the analogously defined share for x-woman

Consider a proxy measure of the number of *eligible* women available to an x-man (in his location at a particular time):

$$\text{eligibles}^m_{x,\ell,t} = \sum_{x'} \mathcal{P}^m_{x,x'} \; w_{x',\ell,t}. \tag{3}$$

Here, we again incorporate the fact that the marriage likelihood between two people depends on their characteristics. Going back to the age example, for a 25-year-old man, a relatively large population of 35-year-old single women does not significantly improve his marriage prospects (in particular, the women "eligible" to him), while a relatively large population of 25-year-old single women does. Unlike the x'-specific competition measure, this *eligibles* measure is not "close to 1."[7] Further, it uses the overall probability ($\mathcal{P}^m_{x,x'}$) instead of the relative probability ($P^m_{x,x'}$). That is, the man's eligibles depend on his likelihood of getting married in the first place (γ^m_x).

Next, define a proxy measure of the number of *competitors* faced by an x-man:

$$\text{competitors}^m_{x,\ell,t} \;=\; \sum_{x'} P^m_{x,x'} \;\text{eligibles}^w_{x',\ell,t} \;=\; \sum_{x'} P^m_{x,x'} \;\sum_{x''} \mathcal{P}^w_{x',x''}\; m_{x'',\ell,t}.$$

For intuition, think of the man as competing in different markets. For example, considering $x = \{age\}$, the man competes in the market for 27-year-old women, 28-year-old women, 29-year-old women, etc. The "competitors" the x-man faces in the market for x'-women are the "eligibles" those women have: $\text{eligibles}^w_{x',\ell,t}$. To get an overall measure of competitors, we weight these eligibles in the x'-market by the man's relative probability of marriage success in that market, $P^m_{x,x'}$.

The overall measure of competition faced by an x-man is simply the competitors they face divided by their eligibles:

$$C^m_{x,\ell,t} \;=\; \text{competitors}^m_{x,\ell,t} \Big/ \text{eligibles}^m_{x,\ell,t}. \tag{4}$$

Just as for the x'-specific measure in (2), this overall measure accounts for certain pairings being more likely than others. Further, it adjusts for the fact that certain groups are more likely to get married. For example, a 30-year-old single woman is more likely to get married in the next year than a 45-year-old single man. To see this, we rewrite the competition

[7]Indeed, in our empirical analysis, we will inform population measures (e.g., the value $W_{x',\ell,t}$) using a sum of observation weights.

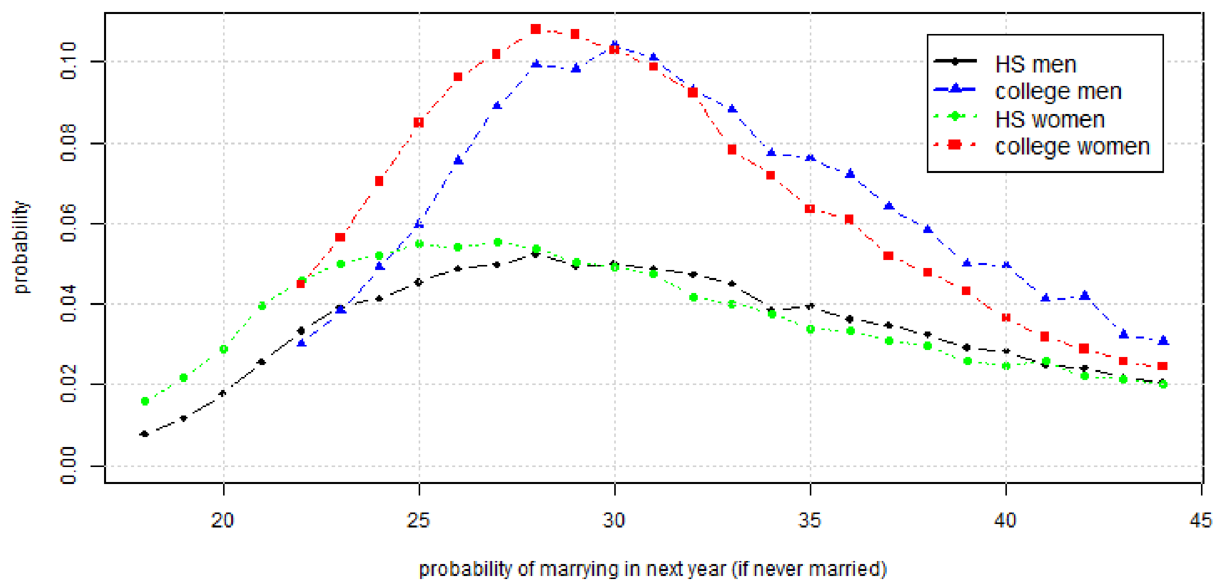


Figure 2: Marriage probabilities by sex and education

Notes: For each sex-education group, we plot marriage probabilities (in the next year) of never-married individuals by age. Probabilities for 18-21-year-old college individuals are omitted because of sparsity.

measure as such:[8]

$$C^m_{x,\ell,t} \;=\; \sum_{x'} P^m_{x,x'}\; w_{x',\ell,t}\left(\frac{\gamma^w_{x'}}{\gamma^m_x}\right) c^m_{x',\ell,t} \Bigg/ \sum_{x'} P^m_{x,x'}\; w_{x',\ell,t}. \tag{5}$$

Here, the competition measure is expressed as the average x'-specific competition measure ($c^m_{x',\ell,t}$) adjusted by the marriage probability ratio ($\gamma^w_{x'}/\gamma^m_x$) and weighted by the share of women in that market ($w_{x',\ell,t}$) times the man's relative probability of success in that market ($P^m_{x,x'}$). Further, one can see that $C^m_{x,\ell,t}$ will be "close to 1" in a similar sense to $c^m_{x,\ell,t}$ (although the former contains a marriage probability ratio while the latter does not). However, as described in Appendix A.3, its average value deviates from 1 notably for different specifications of x. This is unsurprising as the measure is not designed to enforce symmetry.

The formulation in (5) also makes it clear that marriage probabilities (γ^m_x, $\gamma^w_{x'}$) are im-

[8]Again, all relevant objects are defined analogously for women: $c^w_{x',\ell,t}$, eligibles$^m_{x,\ell,t}$, competitors$^m_{x,\ell,t}$, and $C^w_{x,\ell,t}$.

portant for the competition measure. Figure 2 plots marriage probabilities by age (from 18 to 45) for the four sex-education combinations (used in the measure with education, $x = \{age, education\}$). Younger high school women are more likely to get married than high school men, up until age 30, when the likelihoods flip. The same pattern holds for college individuals. So not only are women more likely to get married than men, but they also get married younger. This is in part because women marry men older than them on average (e.g., see Figure 1).

As mentioned, our competition measure offers advantages over the often-used sex ratio. The two main advantages are as follows. First, the typical sex ratio includes individuals regardless of their marriage or cohabitation status. In contrast, our measure only considers "singles" (never married and not cohabitating with an unmarried partner). Second, the sex ratio implicitly assumes that (for example) all heterosexual men are competing with the same group of men for the same group of women. Our measure considers a specific competition environment for an individual based on their characteristics and observed marriage pairings in the data.

2.3 Empirical design

Now consider an outcome $y^s_{x,\ell,t}$ for all single individuals of sex $s \in \{m, w\}$ and characteristics x in location ℓ and year t. To assess the effect of competition on this outcome, we run the following OLS regression:

$$y^s_{x,\ell,t} \; = \; \beta_0 + \beta \; \log(C^s_{x,\ell,t}) + \lambda_{\text{age},\ell} + \tau_{\ell,t} + \epsilon_{x,\ell,t} \tag{6}$$

The main coefficient of interest is β. If we are considering the non-participation rate (for example) and interpreting causally, β is the effect of a 1% increase in the competition measure on the non-participation rate in percentage points. $\lambda_{\text{age},\ell}$ is an age-location fixed effect and

$\tau_{\ell,t}$ is a location-time fixed effect.[9] The purpose of the age-location fixed effect is to net out time-invariant differences in the age-location-specific behavior. For example, location-specific culture can drive the age at which people typically get married and gender-specific work expectations. The purpose of the location-time fixed effect is to net out time-varying forces that vary by location. For example, gender-biased local labor demand shocks can drive changes in competition and labor market outcomes simultaneously. Decisions about fixed effects are discussed further in Section 4.2. We cluster standard errors at the location level because local shocks (like policy, demographic, labor demand, health, etc.) may affect all ages within a location and persist over time. In all regressions, each cell is weighted by the sum of individual ACS weights that make up the cell.

Using OLS, it is likely that the regressor, $\log(C^s_{x,\ell,t})$, is endogenous and the estimate of β is biased. For one, while we expect competition to affect the labor market outcome variables of interest, the reverse could be true as well. In particular, local population composition likely responds to local labor market outcomes. Also, an unspecified variable could be driving movements in the regressor and outcome simultaneously. For example, a negative labor demand shock could simultaneously increase male non-participation and skew the competition measure (e.g., if women are more likely to migrate out for education).[10]

To address this possible endogeneity problem, we employ an IV approach to complement the standard OLS approach. For an instrument to be valid, it must be correlated with the competition measure, and it must affect the outcome variable only through its effect on the competition measure.

Recall the competition measure relies on time- and location-invariant marriage probabilities, and location-specific populations for each characteristic group in a given year. In what

[9]Note that this implicitly assumes age in one of the characteristics in x, which it is in the two measures we consider.

[10]A few other possible endogeneity threats are worth mentioning. First, in locations with strong norms about marriage, gender roles, and family formation, culture may influence both labor supply decisions and the competition measure (through the single population). Second, health-related shocks can also affect both. For example, the opioid crisis likely affected labor supply decisions and population distributions via migration patterns and premature deaths.

follows, we propose a Bartik-style instrument that approximates the competition measure by approximating the location-specific populations. To do so, we use base-year populations from the location of interest and national-level population growth for each group. Since population patterns are likely to persist over time, at least to some extent, the instrument should be highly predictive of the actual competition measure.

This approach corresponds to a pooled exposure research design (the "shares view" of Goldsmith-Pinkham et al. (2020)): locations with different initial demographic compositions are differentially exposed to common national demographic trends over time. The identifying assumption is that, conditional on fixed effects, initial demographic composition is unrelated to subsequent trends in local labor market outcomes except through its interaction with national demographic changes, which shift subsequent local demographic composition. This structure isolates variation in competition driven by national demographic trends rather than by local shocks, migration, or other idiosyncratic forces. We return to the plausibility of this identifying assumption and provide several supporting diagnostics and robustness checks in Section 4.

To define the instrument, some context is necessary. The growth of x-men in location ℓ and year t relative to a base year (t_0) – which we call $g^m_{x,\ell,t}$ – satisfies

$$M_{x,\ell,t} \;=\; \left(1 + g^m_{x,\ell,t}\right)\, M_{x,\ell,t_0}.$$

Note that this growth, $g^m_{x,\ell,t}$, is the sum of national-level growth, $g^m_{x,t}$, and idiosyncratic growth, $\bar{g}^m_{x,\ell,t}$:

$$g^m_{x,\ell,t} \;=\; g^m_{x,t} \;+\; \bar{g}^m_{x,\ell,t}.$$

Similarly, the growth of the whole population in location ℓ and year t relative to a base year – call it $g_{\ell,t}$ – satisfies

$$pop_{\ell,t} \;=\; \left(1 + g_{\ell,t}\right)\, pop_{\ell,t_0},$$

where $g_{\ell,t}$ is the sum of national-level growth, g_t, and idiosyncratic growth, $\bar{g}_{\ell,t}$:

$$g_{\ell,t} \;=\; g_t \;+\; \bar{g}_{\ell,t}.$$

Define $\tilde{m}_{x,\ell,t}$ to be an approximation of $m_{x,\ell,t}$ (the share of x-men in location ℓ and time t) using base-year populations and national-level growth figures:

$$\tilde{m}_{x,\ell,t} \;=\; \frac{(1+g^m_{x,t})M_{x,\ell,t_0}}{(1+g_t)\; pop_{\ell,t_0}}.$$

Recalling that $m_{x,\ell,t} = M_{x,\ell,t}/pop_{\ell,t}$, we are using $(1+g^m_{x,t})M_{x,\ell,t_0}$ to approximate the numerator, $M_{x,\ell,t}$, and $(1+g_t)\; pop_{\ell,t_0}$ to approximate the denominator, $pop_{\ell,t}$.

With these, we can now define the approximation of $C^m_{x,\ell,t}$ as

$$\mathcal{C}^m_{x,\ell,t} \;=\; \underbrace{\sum_{x'} P^m_{x,x'} \sum_{x''} \mathcal{P}^w_{x',x''}\; \tilde{m}_{x'',\ell,t}}_{\text{approximation of competitors}^m_{x,\ell,t}} \Bigg/ \quad \underbrace{\sum_{x'} \mathcal{P}^m_{x,x'}\; \tilde{w}_{x',\ell,t}}_{\text{approximation of eligibles}^m_{x,\ell,t}} \quad . \tag{7}$$

All objects are again defined analogously for women: $g^w_{x,\ell,t}$, $g^w_{x,t}$, $\bar{g}^w_{x,\ell,t}$, $\tilde{w}_{x,\ell,t}$, and $\mathcal{C}^w_{x,\ell,t}$. The idea is that $\mathcal{C}^s_{x,\ell,t}$ will only affect outcome variables through $C^s_{x,\ell,t}$ but is uncorrelated with the "idiosyncratic component" of $C^s_{x,\ell,t}$ (i.e., $\mathcal{C}^s_{x,\ell,t}$ is uncorrelated with $C^s_{x,\ell,t} - \mathcal{C}^s_{x,\ell,t}$).

Now, consider the following two-stage least squares (2SLS) regression:

$$y^s_{x,\ell,t} \;=\; \beta_0 + \beta\;\log(\hat{C}^s_{x,\ell,t}) + \lambda_{\text{age},\ell} + \tau_{\ell,t} + \epsilon_{x,\ell,t}, \tag{8}$$

$$\log(C^s_{x,\ell,t}) \;=\; \gamma_0 + \gamma\;\log(\mathcal{C}^s_{x,\ell,t}) + \lambda_{\text{age},\ell} + \tau_{\ell,t} + \nu_{x,\ell,t}. \tag{9}$$

In our results, we consider 2SLS estimates alongside the OLS estimates.

2.4 Descriptive analysis of the competition measures

Appendix A.3 offers a descriptive analysis of the competition measures. We summarize that analysis here. Figure A.1a plots the average baseline measure for men and women by age. At each age, men experience higher competition. This is what we would expect since men get married later on average, and so there are relatively more single men than single women at any given age. Averaging over all ages 18-45, the measure is 1.27: 1.54 for men and 1.01 for women. Deviation from unity is expected since the measure is not designed to enforce symmetry. We also plot the proxy measures of competitors (the numerator) and eligibles (the denominator) in Figures A.1b and A.1c, respectively, and offer a further explanation of trends. Then we give analogous plots and explanations of trends for the measure with education in Figure A.2.

Next, following Goldsmith-Pinkham et al. (2020), Tables A.2 and A.3 report cross-sectional regressions from the base year (2005) of competition on covariates. While initially high- and low-competition areas differ on some dimensions, this is not problematic since our empirical strategy relies on within-area changes over time, absorbing these baseline differences via fixed effects. Lastly, Figures A.3 and A.4 offer a heat map of the competition measure and its change over time, respectively.

3 Results

This section details our empirical results. We start with the first stage and some other preliminaries, and then analyze the probability of getting married before turning to our main outcome variables. For 25-39-year-olds, we focus on three outcome variables: the non-participation rate, hours worked, and income (for the latter two, we consider the log). For each combination of outcome and sex, we consider two specifications of the baseline competition measure: one for the high school population and another for the college population.[11]

[11]For a given sex and for each cell, while the cell-level competition measure will be the same for the high school and college specifications, the outcome variable will be different, as it is computed using either the

For 18-24-year-olds, we focus on a single outcome variable, the idleness rate, for both the high school and college population combined. While other combinations of population and outcome variables are considered, we focus on these four in the exposition and robustness analysis. Results for the baseline measure are given in this section. As a supplement, analogous results from the measure with education are given in Appendix D.2 alongside the baseline measure results.[12]

3.1 First-stage results and initial context

For our IV approach to be valid, our instrument – the approximation of the competition measure – must be correlated to the actual competition measure. To evaluate whether this is true, we look at first-stage results. Tables D.1 and D.2 in Appendix D show these for our 25-39 single populations (male and female, respectively), and Table D.3 shows these for our 18-24 single populations.[13] First-stage F and Wald statistics are given in these tables (along with the main results tables).

Across competition measures, education levels, and populations, the coefficient on the approximation is large and highly significant. Further, F and Wald statistics are large. This evidence suggests that the competition approximation is highly predictive of the actual competition measure across the board.

For our IV approach to be necessary relative to OLS, the regressor (i.e., the competition measure) must be endogenous. In our main result tables, we evaluate this for each spec-

cell's high school or college population, respectively.

[12]For the 25-39-year-old outcomes, $y^s_{a,\ell,t}$ is the average outcome for all high school (or college) individuals of sex s and age a in PUMA ℓ and year t; e.g., the average log income of all age-33 college women in a particular location and year. However, the baseline competition measure (when $x = \{age\}$ in (4)) does not vary across education within a sex-age-location-year cell. This changes for the measure with education, which computes separate measures for high school and college populations. For the 18-24-year-old specifications which consider high school and college populations together, the cell-level competition measure with education is the average from the two education groups weighted by the single population.

[13]First-stage results may differ slightly across outcome variables because data subsets make slightly different restrictions. For brevity, among outcome variables for 25-39-year-olds, we only report the first-stage results from the non-participation rate. This is the outcome variable that imposes the fewest data restrictions. For 18-24-year-olds, we only consider one outcome variable (the idleness rate), so we report first-stage results for that variable.

ification using a Wu-Hausman test. For low p-values (e.g., below 0.05), the test provides evidence that the regressor is endogenous, and so the 2SLS estimate is preferred. For higher p-values, the test provides no such evidence (either in favor of or against endogeneity), and so the OLS estimate is preferred.

In our results that follow, OLS estimates are consistently statistically close to zero, while 2SLS estimates can be large and statistically significant. Attenuation bias seems likely since the competition measure is constructed from small and noisy sex-age-location-year (or sex-education-age-location-year) cells that fluctuate idiosyncratically.[14]

3.2 Probability of getting married

In what follows, we examine whether marriage competition affects economic outcomes for single adults. Before we do, it seems relevant to test whether such competition affects marriage itself. Consider the population of ACS observations that are "single last year," i.e., single this year or married in the last year. Define the *marriage probability* to be the fraction of those single last year who became married in the last year. Following the outcome variables, consider this probability at the cell level (i.e., the sex-age-year-location level for the baseline measure).

The results are in Table 1 for both 25-39-year-olds (broken down by education and sex) and 18-24-year-olds (broken down by sex only). For 25-39-year-olds, marriage competition significantly decreases the probability of marriage in all specifications at the 1% level. Estimates are biggest for college men: interpreting causally, for a 1% increase in competition, the marriage probability decreases by 0.30 percentage points (pp), a 95% confidence interval (CI) of -0.36 to -0.25 pp. For 18-24-year-olds, estimates are generally smaller in magnitude.[15]

[14]Since OLS uses all variation in the competition measure to inform estimates, this noise dilutes the relationship between competition and the outcome variable, driving estimates toward zero. In other words, OLS likely suffers from attenuation bias because of classical measurement error in the competition measure. In contrast, our instrument only uses the part of competition that would have occurred if the location experienced no local shocks at all. In particular, the instrument is constructed from base-year populations and national demographic growth, so it is mechanically orthogonal to the idiosyncratic cell-level noise that generates measurement error in the competition measure.

[15]Note that the number of observations differs across the high school and college samples. This is because

education	HS		college	
sex	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	-0.0157***	-0.0123***	-0.0454***	-0.0427***
	(0.0034)	(0.0035)	(0.0055)	(0.0057)
2SLS				
Competition (log)	-0.1062***	-0.0744***	-0.3036***	-0.2621***
	(0.0162)	(0.0122)	(0.0271)	(0.0205)
Fit statistics				
Observations	219,484	191,932	143,558	147,948
F-test (1st stage)	12,345.3	19,824.4	8,456.0	15,671.7
Wald (1st stage)	784.1	1,200.2	582.5	1,199.4
Wu-Hausman, p-value	4.72×10^{-11}	6.2×10^{-8}	5.92×10^{-28}	7.95×10^{-36}

(a) 25-39-year-old single adults

subset	*men*	*women*
	(1)	(2)
OLS		
Competition (log)	-0.0087***	-0.0148***
	(0.0015)	(0.0025)
2SLS		
Competition (log)	-0.0286***	-0.1097***
	(0.0088)	(0.0120)
Fit statistics		
Observations	120,231	119,610
F-test (1st stage), Competition (log)	4,555.9	8,532.6
Wald (1st stage), Competition (log)	708.4	971.3
Wu-Hausman, p-value	0.0062	3.28×10^{-23}

(b) 18-24-year-old single adults

Table 1: Effects of marriage competition on marriage probability

Notes: This table gives OLS and 2SLS results for the effect of marriage market competition on the probability of getting married in the next year. The top panel is for 25-39-year-old single adults broken down by education and sex. The bottom panel is for 18-24-year-old single adults broken down by sex only. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

3.3 Hours worked and income

Next, we turn to the effect of marriage competition on log hours worked (per week) and log income for employed single men and women age 25-39. The estimates are in Table 2. For male hours, the results are polarized by education. For high school men, estimates are *negative*

not every sex-age-location-year combination has both high school and college observations. If they did, the number of observations would be consistent. Also, recall that the competition measure for women, $C^w_{a,\ell,t}$, is distinct from that for men, $C^m_{a,\ell,t}$. While higher values of $C^m_{a,\ell,t}$ indicate that *men* face more competition, higher values of $C^w_{a,\ell,t}$ indicate that *women* face more competition.

education	HS		college	
sex	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	0.0109	0.0135	0.0022	-0.0104
	(0.0092)	(0.0119)	(0.0107)	(0.0103)
2SLS				
Competition (log)	-0.1011**	-0.0081	0.1139**	-0.0083
	(0.0393)	(0.0332)	(0.0458)	(0.0293)
Fit statistics				
Observations	178,640	149,872	122,626	128,069
F-test (1st stage)	10,121.1	16,454.4	7,417.7	14,524.9
Wald (1st stage)	687.4	1,076.8	493.9	1,105.4
Wu-Hausman, p-value	0.0010	0.4855	0.0048	0.9397

(a) log hours

competition measure	**baseline**			
education	HS		college	
sex	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	0.0185	0.0369	0.0382	-0.0126
	(0.0239)	(0.0289)	(0.0347)	(0.0300)
2SLS				
Competition (log)	-0.0297	-0.1027	1.119***	0.1284
	(0.1189)	(0.1036)	(0.1494)	(0.0919)
Fit statistics				
Observations	178,640	149,872	122,626	128,069
F-test (1st stage)	10,121.1	16,454.4	7,417.7	14,524.9
Wald (1st stage)	687.4	1,076.8	493.9	1,105.4
Wu-Hausman, p-value	0.6012	0.0776	4.73×10^{-20}	0.0516

(b) log income

Table 2: Effects of marriage competition: hours worked and income

Notes: These two panels have the same structure as Table 3a and are for the same population (i.e., single adults not in school age 25–39), but detail results for log hours worked and log income, respectively. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

and significant at the 5% level: a 1% increase in competition decreases hours worked by 0.10%. For college men, estimates are *positive* and significant at the 5% level: a 1% increase in competition increases hours worked by 0.12%. This suggests an interesting pattern: men with a college education (and potentially better labor market prospects) appear to respond to tighter marriage markets by increasing their labor supply on the intensive margin while the opposite holds for their non-college-educated counterparts. For women, on the other hand, the estimates are not statistically distinguishable from zero.

For college male income, the 2SLS estimate is substantial and statistically significant: a 1% increase in competition increases income by 1.12% (95% CI: 0.82 to 1.42%). The estimates for high school men and both high school and college women are statistically close to zero. This seems to suggest that while the increase in the labor supply by college-educated men translates into higher earnings, the opposite does not appear to happen for high school men, and we observe no effect on the extensive margin.[16]

3.4 Non-participation and idleness

Table 3a shows non-participation rate results for single adults not in school aged 25-39. For men, the 2SLS estimate in the first column suggests that a 1% increase in competition decreases non-participation for high school males by 0.10 pp (95% CI: -0.19 to -0.01 pp). The college male estimate (third column) is slightly smaller in magnitude and is only statistically significant at the 10% level.

The second and fourth columns detail the analogous estimates for women (single, not in school, age 25-39). In contrast to the corresponding male estimate, the female estimates are mostly positive. For high school women, the 2SLS estimate for the baseline measure is 0.11 pp, which is significant at the 95% level (95% CI: 0.02 to 0.20 pp). The OLS estimates are also positive, but relatively small and only significant at the 10% level.

Next, we look at the idleness rate – the fraction of individuals not working and not in school – among single 18-24-year-olds. Table 3b details the results. The estimate for men is significantly negative: for a 1% increase in competition, the idleness rate decreases by 0.11 pp (95% CI: -0.15 to -0.06 pp). For women, the corresponding estimate is small and

[16]As a complement, Table D.7 in Appendix D gives the analogous results to Table 2, but for 18-24-year-olds (single, not in school). For these younger women, both hours and income significantly increase. For male income, the estimate is negative. For male hours, the Wu-Hausman value is high, so the OLS estimate is preferred, and it is statistically close to zero. A caveat to these complementary results is that analyzing income for 18-24-year-olds is potentially problematic. The income variable in the ACS uses income from the last 12 months. If someone finished school and started a job a few months before their ACS interview, their last 12 months of income would not be representative of their annual income moving forward (to our knowledge, there is no way to remove those recently in school to address this issue). Those 18-24 are more likely to have just finished school, which is why our main income analysis focuses on those 25-39. Hours may suffer from the same issue since it also uses the previous 12 months as the reference period.

education	HS		college	
sex	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	-0.0049	0.0238*	0.0060	0.0154*
	(0.0108)	(0.0135)	(0.0102)	(0.0085)
2SLS				
Competition (log)	-0.0991**	0.1076**	-0.0803*	0.0178
	(0.0455)	(0.0465)	(0.0415)	(0.0273)
Fit statistics				
Observations	215,378	183,229	133,564	135,788
F-test (1st stage)	12,161.4	19,491.5	7,931.8	15,380.9
Wald (1st stage)	765.7	1,174.9	534.5	1,145.5
Wu-Hausman, p-value	0.0194	0.0233	0.0131	0.9156

(a) Non-participation rate for 25-39-year-old single adults

competition measure	**baseline**	
subset	*men*	*women*
	(1)	(2)
OLS		
Competition (log)	-0.0015	-0.0077
	(0.0041)	(0.0058)
2SLS		
Competition (log)	-0.1074***	-0.0165
	(0.0237)	(0.0216)
Fit statistics		
Observations	120,231	119,610
F-test (1st stage)	4,555.9	8,532.6
Wald (1st stage)	708.4	971.3
Wu-Hausman, p-value	1.69×10^{-7}	0.6596

(b) idleness rate for 18-24-year-old single adults

Table 3: Effects of marriage competition: non-participation and idleness

Notes: This table gives OLS and 2SLS results for two extensive margin outcomes. The top panel is the non-participation rate (i.e., the non-participation rate) for single adults not in school, aged 25–39, broken down by education and sex. The bottom panel is the idleness rate (i.e., the fraction of individuals not working and not in school) among single 18–24-year-olds, broken down by sex. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

statistically close to zero. As a complement, Table D.6a in Appendix D gives the analogous results for idleness among 25-39-year-old single adults. Patterns in these estimates match those for non-participation in Table 3a. Further, Table D.6b gives the analogous results for college attendance (undergraduate or graduate school) among 18-24-year-olds. For men, competition significantly increases college attendance, suggesting that the drop in idleness is driven by a rise in college attendance. For women, the estimate is also significantly positive.

4 Robustness and Extensions

This section offers a discussion of important methodological decisions and robustness. The complete robustness analysis and a few extensions are detailed in Appendix B, and all referenced tables and figures are in Appendix D.

4.1 Competition measure

Up to this point, we have only considered the baseline competition measure, which uses age in the vector of characteristics. As mentioned, Appendix D.2 repeats the main analysis with both age and education in this vector (the *measure with education*). For most education-sex-outcome combinations, estimates are close across measures. However, a few specifications, estimates from the two competition measures differ meaningfully.[17] For completeness, in the robustness analysis and extensions that follow, we include results for the measure with education alongside those from the baseline measure, although we only focus on the latter in our discussion.

4.2 Fixed effects

Recall our identifying assumption: conditional on the included fixed effects, the *initial* sex–age composition of a PUMA is unrelated to subsequent local labor market outcomes except through its interaction with national demographic trends (which differentially shift future local sex–age composition across locations). In this subsection, we argue that our particular choice of fixed effects helps make this assumption plausible.

Age-location fixed effects remove permanent differences across places and ages; e.g., they net out any time-invariant differences between the behavior of 30-year-old men in Charleston and that of 30-year-old men in Los Angeles. In particular, cultural differences can drive

[17] For example, for 18-24-year-old women, the idleness estimate is significantly negative for the measure with education (-0.19 pp; 95% CI: -0.21 to -0.18) but statistically close to zero for the baseline measure (-0.01 pp; 95% CI: -0.02 to 0.00; note that OLS is preferred).

marriage activity and labor market behavior that vary across sex and age within a location.[18] Location-time fixed effects remove the impact of local shocks; e.g., they absorb any shock in Orlando in 2015 that affects all ages.[19]

The variation that remains is within an age-location cell over time. It comes from differential *exposure* to national demographic trends, which is dictated by initial populations. The larger an initial sex-age population in a particular location, the greater the exposure to national-level population trends for that sex-age. For example, the larger the initial population of 32-year-old women in Charlotte, the more exposed Charlotte is to the national growth of 32-year-old women. Since the identifying variation comes from differences across age-location cells over time, age-time fixed effects would be inappropriate.[20]

Table D.8 in Appendix D considers different fixed-effect combinations for the high school men non-participation rate regression. These results highlight the importance of the age-location fixed effect. Further, the instrument is strong across specifications, and there is consistency between estimates with age-location and location-time fixed effects, and age-location and time fixed effects. Figure D.1 in Appendix D plots the residuals from several of these specifications to gauge whether adding certain fixed effects eliminates too much variation. While residual variation visually decreases as fixed effects are added, there is still notable variation in the most saturated model – i.e., that with age-location and location-time

[18]For example, in certain areas, women will be expected to marry younger and have less labor market attachment, while in other areas, women are less likely to marry altogether and have strong labor market attachment. These cultural differences will impact the competition measure and outcome variables, such as non-participation and hours worked. Note that separate age and location fixed effects do not capture this possible trend, whereas age-location fixed effect do.

[19]Such shocks can take different forms. For instance, gender-biased local labor demand shocks can drive changes in competition and labor market outcomes simultaneously (e.g., men sorting into resource-extractive jobs like mining, increasing their relative wages and the marriage competition they face simultaneously). Another example is the opioid epidemic. Opioid exposure timing and intensity varied over locations, which likely affected labor market outcomes and the marriage market.

[20]Since the instrument uses national age-time-specific information to predict age-location-specific populations, adding age-year fixed effects would take much of the predictive power from the instrument in the first stage. For example, age-time fixed effects would absorb any national change that is common to all 26-year-old women. But the instrument relies on the national growth of 26-year-old women. As expected, when we include them (results unreported), first-stage coefficients weaken, and 2SLS estimate standard errors increase substantially. In summary, age-time fixed effects would absorb the national demographic trends that generate differential exposure across locations, eliminating the variation used for identification.

fixed effects, which we use throughout our analysis.

4.3 Instrument validity

To understand what would invalidate our instrument, consider what must be true for the identifying assumption to fail. The instrument relies on initial sex–age composition to generate differential exposure to national demographic trends. It would therefore be invalid if places with different initial demographic compositions were already on systematically different trajectories of labor market outcomes for reasons unrelated to marriage market competition.[21] Put differently, the instrument fails if initial demographic structure proxies for latent forces that directly drive future labor market outcomes. Our fixed-effects structure is designed to mitigate these concerns: age–location fixed effects absorb time-invariant differences across places and ages, while location–time fixed effects absorb shocks that vary across locations over time. The remaining variation comes from differences in how locations are exposed to common national demographic trends through their initial demographic composition.

In many environments with Bartik-style instruments, Rotemberg weights provide useful diagnostics. They can allow researchers to see which industries, groups, shocks, etc., are actually driving identification and to assess whether the design is relying heavily on a small or potentially problematic subset of variation. However, they seem inappropriate in our setting.[22]

Instead of Rotemberg weight diagnostics, we assess the identifying assumption by ex-

[21] For example, locations with relatively many 33-year-old women in the initial year might also be places undergoing cultural shifts away from female labor force participation and toward family specialization. Similarly, locations with relatively many 29-year-old men in the initial year might be places experiencing declines in male labor demand.

[22] In the canonical example of Goldsmith-Pinkham et al. (2020), which considers a *linear* shift-share instrument, the 2SLS estimator can be decomposed in a straightforward way using these weights. However, in our setting, the competition measure (and its instrument) is a *nonlinear* function of demographic cell counts with probability weights that enter both the numerator and denominator (and, in our baseline specification, through a log transformation). Because our instrument is more complicated, the construction of Rotemberg weights or something similar would not be straightforward. Even if they were constructed, they would be less interpretable and, thus, likely less useful.

amining whether the instrument predicts pre-existing trends in labor market outcomes. In Appendix B.1, we consider a few relevant tests (see Tables D.9, D.10, and D.11), most significantly, an augmented specification with age–location–specific linear time trends, which allow each age group within each PUMA to follow its own gradual trajectory over time. Identification is, thus, restricted to shorter-run deviations from those trends, and standard errors are notably higher. Because demographic composition evolves gradually over time, this specification is particularly demanding and absorbs much of the variation used to identify the baseline effects.

Despite the loss of precision, the non-participation effect for high school women (25-39) and the idleness effect for men (18-24) remain significant. This suggests that these effects are not driven solely by slow demographic shifts in local marriage markets. In contrast, the college male hours and income estimates and the high school male hours and non-participation estimates lose significance. This suggests that the corresponding estimates from the main analysis may rely more heavily on longer-run demographic variation in local marriage markets. However, given the substantial reduction in identifying variation, these results should be interpreted as inconclusive rather than as evidence against the baseline findings or a violation of the identifying assumption.

4.4 Other checks and extensions

In our main analysis, we use "consistent PUMAs" as the geographical unit since that is the most granular geography offered in the ACS that is consistent over the time period. However, a less granular geography may be more appropriate. For this reason, we consider two alternate ways to divide locations that are generally less granular than PUMAs: commuting zones (CZs) from USDA ERS (2026) and "spillover-defined local labor markets" (SLMs) from Bartik (2024). Results (in Tables D.12 and D.13, respectively) are very close to those from Section 3. In particular, specifications with significant estimates in Section 3 also have significant estimates here.

Recall that our instrument uses an approximation of local populations using national-level growth. Mechanical endogeneity is a common concern for Bartik-style instruments like this; in our case, the location's population could influence national figures. Since we want the instrument to reflect exogenous national trends and not local shocks, this is potentially problematic. Because we use many locations (1,078 PUMAs) mostly around the same size, such mechanical endogeneity is less of a concern. To be safe, we recompute the results using a leave-one-out (LOO) Bartik construction. The results (in Table D.14) are again very close to those from Section 3, which gives us confidence that mechanical endogeneity plays a negligible role.

A potential problem or bias is that the competition measure is not forward-looking. For an age-a single man (woman), the measure uses the age distribution of marriages for an age-a man (woman). In reality, if the person is actually "single" – never married and no boyfriend or girlfriend – and looking for a marriage partner, they are likely thinking about marriage a few years in the future rather than the current year (at least on average). To account for this, we adjust the competition measure by replacing $\mathcal{P}^s_{a,a'}$ and $P^s_{a,a'}$ with $\mathcal{P}^s_{a+c,a'+c}$ and $P^s_{a+c,a'+c}$, respectively, where c stands for "courtship" and c is taken to be 3. The recomputed results are in Table D.15. Estimates are again close to those in Section 3.

For further robustness, we consider a "long difference" approach. Instead of using the log competition value as the endogenous regressor (and the log of the approximated competition value as the instrument), we use the change in competition from a base year (and the change in the approximated competition). The results are reported in Table D.16. Estimates are generally similar to those in the main analysis.

In Appendix B.3, we further consider a few extensions. These specifications are not directly relevant to robustness of the main results because they use different populations or approaches. First, we consider regressions with "competitors" and "eligibles" separated; recall competition is the ratio of competitors to eligibles. Generally, when there is a significantly positive (negative) coefficient on competition in the main results, there is a significantly positive (negative) coefficient on competitors and a significantly negative (positive) coefficient on eligibles, as expected (see Tables D.17 and D.18). Second, to assess whether the main results are stable across distinct economic and social environments, we split the sample in half by time (2006-2013 and 2014-2021) and rerun the main analysis in Tables D.19 and D.20. Several findings remain consistent across both periods, while others vary across periods (suggesting the effect may be concentrated during a specific time).[23] Third, we use "unmarried" individuals – not currently married, but perhaps married before – instead of singles (in both the competition measure construction and the regressions). The results in Table D.21 are generally consistent with the main results, but statistical significance for the high school male hours estimates disappears. Fourth, we use "childless" individuals: singles without any children in the household. For the results in Table D.22, significance holds up for college male income and hours and male idleness, but is attenuated for high school male and female non-participation, and high school male hours. Lastly, we run a specification that computes all model objects using only white observations. Not only are outcomes and populations computed using only whites, but marriage probability distributions are only computed using newlywed whites who marry other whites. In Table D.23, estimates are again consistent with the main results, but significance is attenuated for high school male non-participation and hours, and high school female non-participation.

4.5 Summary and interpretation

Overall, it appears that local marriage markets have sizable effects on labor market outcomes. This is especially the case for college men (single and not in school between age 25-39). For a 10% rise in competition, their hours increase by 1.1% on average and their income increases by 11.2%. For high school men – also single and not in school between ages 25-39 – the picture is different. On one hand, competition appears to decrease non-participation. For

[23]Nevertheless, splitting the sample substantially reduces precision, so the loss of significance in some specifications should not necessarily be interpreted as evidence of no effect. Further, because 2SLS estimators are nonlinear covariance-ratio estimators, pooled estimates are not mechanically equal to averages of split-sample estimates.

a 10% rise in competition, the non-participation rate decreases by 1.0 pp (although this estimate does not remain significantly negative at the 5% level for two of the robustness tests). On the other hand, competition also appears to decrease hours worked. For a 10% rise in competition, hours worked decrease by 1.0%. A simultaneous increase in the extensive margin and decrease in the intensive margin suggests a compositional change could be at play. Also, for younger men (age 18-24), competition appears to decrease idleness. For a 10% rise in competition, the idleness rate decreases by 1.1 pp. Finally, for a 10% rise in competition, high school female non-participation increases by 1.1 pp.

In general, these estimates are robust to a series of tests. See Appendix B.1 for details and Appendix B.2 for a brief summary. One exception is that when we add time trends to the age-location fixed effects, some of these estimates are no longer significant. However, this does not nullify the results since much of the identifying variation is removed. Two specifications that remain significant after adding these time trends are the high school female non-participation increase and the male idleness decrease.

Several possible interpretations of these results exist. However, drawing conclusions may be speculative. For the interested reader, we have detailed a few possible interpretations and connected them to relevant literature in Appendix C.

5 Conclusion

Marriage-market incentives are a potentially overlooked determinant of labor market outcomes. In this study, we investigate this possibility. In particular, we develop a novel tightness measures for local marriage markets that use national data on marriage patterns of 18-45-year-olds and local population data. This measure offers advantages over the typically-used sex ratio. We confirm that higher competition is causally associated with lower marriage probabilities.

Among singles aged 25-39, we find that college-educated men appear to increase their hours worked in response to increased competition, while the opposite holds for men with only a high school education. The former group also increases their earnings. On the extensive margin, non-college-educated men are less likely to be out of the labor force in response to tighter competition while the opposite appears to hold for women. Further, younger men (aged 18-24) reduce their idleness. With a few caveats, these results are robust to alternative specifications.

Recall this paper was motivated by increasing non-participation of young men. While we find a causal link between marriage market competition and non-participation (and idleness), the competition measure does not change much over our analysis period. In Figure D.2 in Appendix D, we plot the two measures (baseline and with education) over time for each sex-education combination. All changes over the period are marginal, suggesting that changes in competition are not driving non-participation trends.

Nevertheless, we find that changes in competition exert nontrivial and heterogeneous effects on individual labor supply and related outcomes. These outcomes suggest potentially significant economic implications of changes in the structure of marriage and dating markets. This is particularly important in the context of decreasing marriage rates and engagement with dating, especially for young men.

References

Ali Akbari, Danial, Thomas Fischer, and Demid Getik (2025). “Is the Child Penalty Sustaining the Glass Ceiling Effect?” In: URL: https://ssrn.com/abstract=5174411.

Amuedo-Dorantes, Catalina and Shoshana Grossbard (2007). “Cohort-level sex ratio effects on women’s labor force participation”. In: *Review of Economics of the Household* 5.3, pp. 249–278.

Angrist, Josh (2002). “How do sex ratios affect marriage and labor markets? Evidence from America’s second generation”. In: *The Quarterly Journal of Economics* 117.3, pp. 997–1038.

Autor, David, David Dorn, and Gordon Hanson (2019). “When work disappears: Manufacturing decline and the falling marriage market value of young men”. In: *American Economic Review: Insights* 1.2, pp. 161–178.

Bartik, Timothy J (2024). *Local labor markets should be redefined: New definitions based on estimated demand-shock spillovers*. Tech. rep. Upjohn Institute Working Paper.

Becker, Gary S (1973). “A theory of marriage: Part I”. In: *Journal of Political Economy* 81.4, pp. 813–846.

— (1974). “A theory of marriage: Part II”. In: *Journal of Political Economy* 82.2, Part 2, S11–S26.

Bertrand, Marianne, Emir Kamenica, and Jessica Pan (2015). “Gender identity and relative income within households”. In: *The Quarterly Journal of Economics* 130.2, pp. 571–614.

Blundell, Richard, Luigi Pistaferri, and Itay Saporta-Eksten (2016). “Consumption inequality and family labor supply”. In: *American Economic Review* 106.2, pp. 387–435.

Bursztyn, Leonardo, Thomas Fujiwara, and Amanda Pallais (2017). “‘Acting wife’: Marriage market incentives and labor market investments”. In: *American Economic Review* 107.11, pp. 3288–3319.

Cesarini, David et al. (2023). *Fortunate families? The effects of wealth on marriage and fertility*. Tech. rep. National Bureau of Economic Research.

Chiappori, Pierre-Andre, Bernard Fortin, and Guy Lacroix (2002). “Marriage market, divorce legislation, and household labor supply”. In: *Journal of political Economy* 110.1, pp. 37–72.

Chiappori, Pierre-André, Monica Costa Dias, and Costas Meghir (2018). “The marriage market, labor supply, and education choice”. In: *Journal of Political Economy* 126.S1, S26–S72.

de Linde Leonard, Megan and T.D. Stanley (2015). “Married with children: What remains when observable biases are removed from the reported male marriage wage premium”. In: *Labour Economics* 33, pp. 72–80. ISSN: 0927-5371. DOI: https://doi.org/10.1016/j.labeco.2015.02.010. URL: https://www.sciencedirect.com/science/article/pii/S0927537115000251.

Getik, Demid (2024). “Relative Income and Mental Health in Couples”. In: *The Economic Journal* 134.664, pp. 3291–3305. ISSN: 0013-0133. DOI: 10.1093/ej/ueae071. eprint: https://academic.oup.com/ej/article-pdf/134/664/3291/60665174/ueae071.pdf. URL: https://doi.org/10.1093/ej/ueae071.

Giuntella, Osea, Lorenzo Rotunno, and Luca Stella (Dec. 2022). “Globalization, Fertility, and Marital Behavior in a Lowest-Low Fertility Setting”. In: *Demography* 59.6, pp. 2135–

2159. DOI: 10.1215/00703370-10275366. eprint: https://read.dukeupress.edu/demography/article-pdf/59/6/2135/1682106/2135giuntella.pdf. URL: https://doi.org/10.1215/00703370-10275366.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift (2020). "Bartik instruments: What, when, why, and how". In: *American Economic Review* 110.8, pp. 2586–2624.

Grossbard-Shechtman, Shoshana A and Shoshana Neuman (1988). "Women's labor supply and marital choice". In: *Journal of Political Economy* 96.6, pp. 1294–1302.

Hitsch, Günter J, Al Hortaçsu, and Dan Ariely (2010). "Matching and sorting in online dating". In: *American Economic Review* 100.1, pp. 130–163.

Johnston, David W., Rachel J. Knott, and Nidhiya Menon (2023). "Violations of the Male Breadwinner Norm and Mental Health". In: Preprint, version dated September 29, 2023. URL: https://ssrn.com/abstract=4617812.

Kearney, Melissa S. and Riley Wilson (Oct. 2018). "Male Earnings, Marriageable Men, and Nonmarital Fertility: Evidence from the Fracking Boom". In: *The Review of Economics and Statistics* 100.4, pp. 678–690. ISSN: 0034-6535. DOI: 10.1162/rest_a_00739. eprint: https://direct.mit.edu/rest/article-pdf/100/4/678/1918770/rest_a_00739.pdf. URL: https://doi.org/10.1162/rest_a_00739.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard (Oct. 2019). "Children and Gender Inequality: Evidence from Denmark". In: *American Economic Journal: Applied Economics* 11.4, pp. 181–209. DOI: 10.1257/app.20180010. URL: https://www.aeaweb.org/articles?id=10.1257/app.20180010.

Korenman, Sanders and David Neumark (1991). "Does Marriage Really Make Men More Productive?" In: *Journal of Human Resources* 26.2, pp. 282–307. DOI: 10.2307/145924. URL: https://doi.org/10.2307/145924.

Lino, Mark et al. (Jan. 2017). *Expenditures on Children by Families, 2015*. Miscellaneous Publication No. 1528-2015, U.S. Department of Agriculture, Center for Nutrition Policy and Promotion. Revised March 2017; available at https://www.fns.usda.gov/cnpp/expenditures-children-families.

Ludwig, Volker and Josef Brüderl (2018). "Is There a Male Marital Wage Premium? New Evidence from the United States". In: *American Sociological Review* 83.4, pp. 744–770. DOI: 10.1177/0003122418784909. URL: https://doi.org/10.1177/0003122418784909.

Lundborg, Petter, Erik Plug, and Astrid Würtz Rasmussen (June 2017). "Can Women Have Children and a Career? IV Evidence from IVF Treatments". In: *American Economic Review* 107.6, pp. 1611–37. DOI: 10.1257/aer.20141467. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20141467.

— (2024). *Is there really a child penalty in the long run? New evidence from IVF treatments*. Tech. rep. IZA Discussion Papers.

Macunovich, Diane J (2012). "Relative cohort size, relative income, and married women's labor force participation: United States, 1968–2010". In: *Population and Development Review* 38.4, pp. 631–648.

McDonald, Patrick (2020). "The Male Marriage Premium: Selection, Productivity, or Employer Preferences?" In: *Journal of Marriage and Family* 82.5, pp. 1553–1570. DOI: 10.1111/jomf.12683.

Missouri Census Data Center (2018). *Geocorr 2018: Geographic Correspondence Engine*. http://mcdc.missouri.edu/applications/geocorr2018.html. Accessed: 2026-03-25.

Pew Research Center (Feb. 2023). *For Valentine's Day, 5 facts about single Americans.* https://www.pewresearch.org/short-reads/2023/02/08/for-valentines-day-5-facts-about-single-americans/. Short reads, February 8 2023 (accessed October 29, 2025).

Pierce, Lamar, Michael S. Dahl, and Jimmi Nielsen (Mar. 2013). "In Sickness and in Wealth: Psychological and Sexual Costs of Income Comparison in Marriage". In: *Personality and Social Psychology Bulletin* 39 (3), pp. 359–374. ISSN: 01461672. DOI: 10.1177/0146167212475321.

Rapoport, Benoıt, Catherine Sofer, and Anne Solaz (2011). "Household production in a collective model: some new results". In: *Journal of Population Economics* 24.1, pp. 23–45.

Ruggles, Steven et al. (2025). *IPUMS USA: American Community Survey (ACS).* Tech. rep. Version 16.0. ACS / IPUMS USA. URL: https://usa.ipums.org/usa/.

Schechtl, Manuel and Nicole Kapelle (2024). "The Male Marital Earnings Premium Contextualized: Longitudinal Evidence from the United States, Germany, and the United Kingdom". In: *Journal of Marriage and Family* 86.1, pp. 176–198. DOI: 10.1111/jomf.12937.

Shenhav, Na'ama (2021). "Lowering standards to wed? Spouse quality, marriage, and labor market responses to the gender wage gap". In: *Review of Economics and Statistics* 103.2, pp. 265–279.

U.S. Department of Agriculture, Economic Research Service (2026). *Commuting Zones and Labor Market Areas: 2020 County-to-Commuting Zone Crosswalk.* https://www.ers.usda.gov/data-products/commuting-zones-and-labor-market-areas. Accessed: 2026-03-25.

Appendices

A Data Appendix

A.1 Justifying our two measures

When choosing specifications for x, there are two primary concerns. On one hand, we want x to capture characteristics relevant to absolute and conditional marriage probabilities (i.e., γ_x^s and $P_{x,x'}^s$, respectively). On the other hand, our data is limited, so a rich specification of x could lead to sparsity problems when computing $P_{x,x'}^s$.

A first variable that is crucial to these probabilities is age. This importance is made clear by Figures 1 and 2 (in the body); the former shows the conditional marriage probabilities of 30-year-old men and women, and the latter shows how the absolute marriage probability varies by age in our sample. In a *baseline measure*, we consider x to include age alone.

A second variable is education. To demonstrate its importance, consider Table A.1, which reports marriage probabilities by education and sex for never-married 25-39-year-olds in the

		high school		**college**	
		men	*women*	*men*	*women*
prob. of marriage in next year by spouse's educ.	*overall*	4.5%	4.7%	8%	8.7%
	high school	3.3%	3.9%	1.6%	2.9%
	college	1.2%	0.8%	6.4%	5.9%
cond. prob. of spouse's educ.	*high school*	73.1%	82.9%	20.1%	32.6%
	college	26.9%	17.1%	79.9%	67.4%

Table A.1: Marriage probabilities by education and sex (25-39-year-olds)

Notes: The first row shows the overall probability of a never-married individual becoming married to anyone in the next year. For example, the entry in the first column indicates that a high school man (never-married, age 25-39) has a 4.5% probability of becoming married in the next year. The second and third rows split this percentage based on the education of their marriage partner: high school or college, respectively. The fourth and fifth columns report conditional probabilities. For example, the first entry of the fourth column indicates that a newly married high school man marries a high school woman 73.1% of the time. “High school” and “college” refer to less than and at least a bachelor’s degree, respectively. Note that homosexual marriages are excluded from these statistics. Data is from the ACS.

ACS. While high school individuals have around a 4-5% chance of being married in the next year, college individuals have an 8-9% chance. Here, “high school” and “college” refer to less than and at least a bachelor’s degree, respectively. Also, note that women are more likely to get married than men. Furthermore, individuals are more likely to marry partners in their own education group. The rest of Table A.1 shows this clearly. For example, high school men marry high school women 73.1% of the time, and college men marry college women 79.9% of the time. To account for these differences, we consider a supplementary specification where x includes both sex and education, the *measure with education*.

A.2 Construction of the competition measure

As mentioned, we use two specifications of x in the analysis: a baseline specification with age only, and a second with age and education. In what follows, we only describe the construction and application of the baseline measure for brevity, replacing the subscript x (for vector of characteristics) with a (for age). The construction of the second measure is analogous.

We build the measure at the PUMA level over the age range 18-45. There are three types of objects used to build measure: (1) single populations at the age-location-year level ($M_{a,l,t}$ and $W_{a,l,t}$); (2) marriage probabilities by age (γ_a^m and γ_a^w); (3) conditional marriage probability distributions at each age ($P_{a,\cdot}^m$ and $P_{a,\cdot}^w$). The first is taken to be the sum of person weights for single men and women in each age-location-year cell.[1] We consider someone as “single” if they have never been married and are not cohabitating with an unmarried partner. The second is the fraction of all observations (of a particular age and sex from the entire data sample) who were never married the previous year, but became married for the first time. This is isolated using variables for whether an individual was married in the last 12 months and how many times they have been married. The computation of the third is more involved.

[1]This is appropriate since the ACS (with the appropriate weights) is designed to be representative of the U.S. population. Also note that we use these populations to compute the national growth figures, $g_{a,t}^m$ and $g_{a,t}^w$, used in the competition measure approximation.

We consider the set of *newlyweds* to be observations: (1) that married in the last 12 months; (2) for which we have information on their spouses' sex, age, and education; (3) that have a spouse of the opposite sex. For each age a between 18 and 45, $P^m_{a,a'}$ is the fraction of newlywed age-a men who married an age-a' woman (before adjustment). Similarly, $P^w_{a,a'}$ is the fraction of newlywed age-a women who married an age-a' man. We adjust raw densities from the data by rescaling for availability. The idea is that the raw probabilities are dependent on the age distributions of available (i.e., single) men and women.[2] For example, in a counterfactual with more single older men, there would be more marriages of younger women to older men (ceteris paribus).

As an illustration of these distributions, consider the distributions for 30-year-old men and women in Figure 1 (in the body). While both distributions have a mode of 30, that for men is left-skewed and that for women is right-skewed. For example, a 30-year-old woman is just as likely to marry a 40-year-old man as a 25-year-old man, while a 30-year-old man is much more likely to marry a 25-year-old woman than a 40-year-old woman. In summary, a 30-year-old newlywed man is more likely to have a younger wife, and a 30-year-old newlywed woman is more likely to have an older husband.

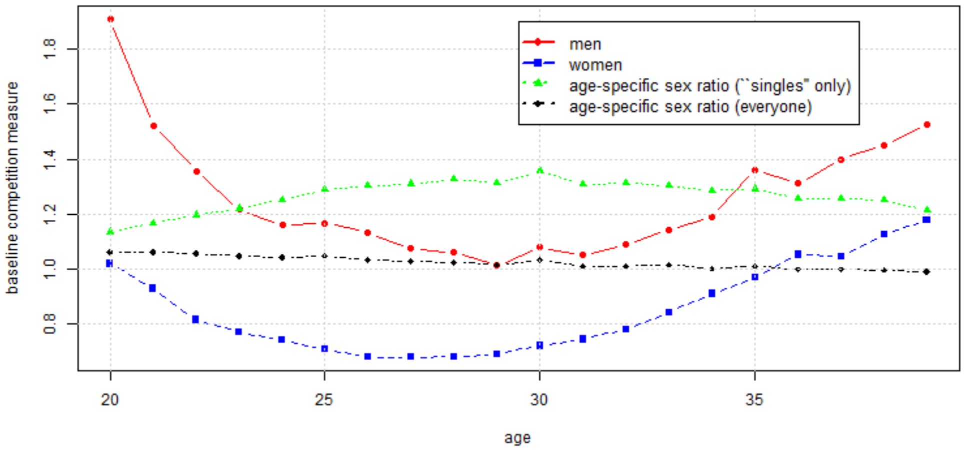
A.3 Descriptive analysis of the competition measures

Because men get married later, there are relatively more single men than single women at a given age. Naturally, we would expect men to experience more competition (as we measure it) than women. This is what we find in Figure A.1a. Here, we plot the average baseline

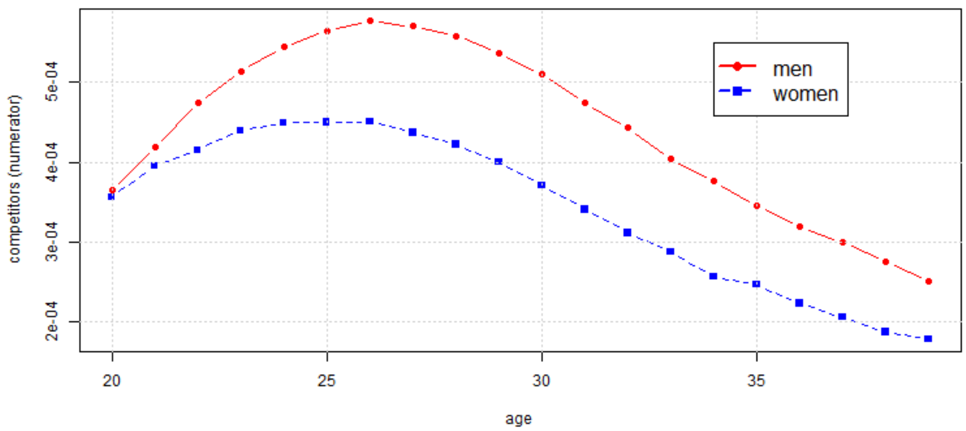
[2]When rescaling, assume that if you change the number of single women (men) of age a', this will increase the relative probability of an age a man (woman) marrying an age a' woman (man) by approximately the same amount. In particular, assume $\{q^s_{a,a'}\}_{s,a,a'}$ are the raw probabilities for each sex s and ages a and a'. The rescaled probabilities are $\{Q^s_{a,a'}\}_{s,a,a'}$ which satisfy

$$Q^s_{a,a'} = (q^s_{a,a'} \;/\; S^{s'}_{a'}) \Big/ \sum_{a''} (q^s_{a,a''} \;/\; S^{s'}_{a''}),$$

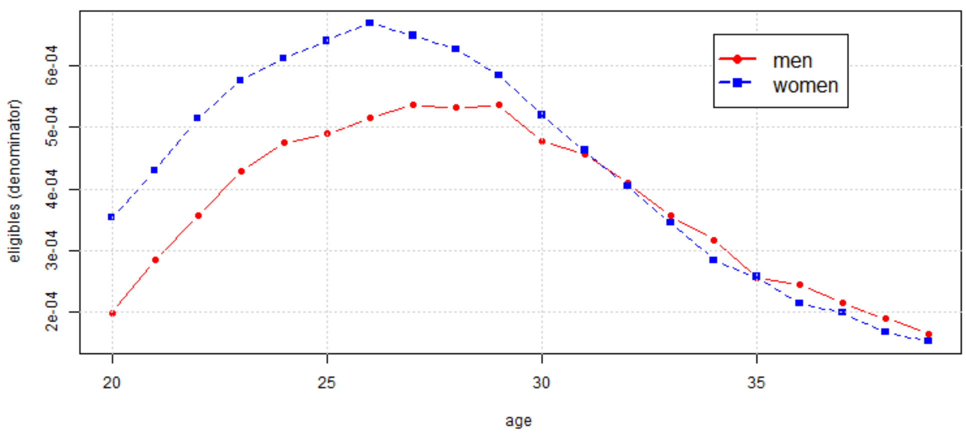
and $S^{s'}_{a'}$ is the national level fraction of age a' individuals of the opposite sex ($s' \neq s$) who are single. For example, if 1/3 of age-35 women are single, $S^w_{35} = 1/3$, then the probability a newlywed age-30 man marries an age-35 woman increases by around 3 times. But this does not account for the denominator, which is necessarily greater than 1, so the increase will be by less than a factor of 3.



(a) Baseline competition measure



(b) Competitors (numerator)



(c) Eligibles (denominator)

Figure A.1: Baseline competition measure, competitors, and eligibles by sex and age

Notes: See equation (4) for the relevant breakdown: competition = competitors / eligibles. The age-specific sex ratio for singles only (i.e., never married and not cohabitating with an unmarried partner) and all individuals are given for reference.

competition measure for men and women by age. At every age, the average value for men is well above that for women. Alongside the competition measure, we also plot the proxy measures of competitors (the numerator) and eligibles (the denominator) in Figures A.1b and A.1c, respectively. The age-specific sex ratios (of singles and all individuals) are also plotted in Figure A.1a for reference.

In addition to the sex disparity, several points are worth noting. First, the age profile of eligibles is an upside-down-U shape with a long right tail. This roughly mirrors the marriage probability age-profile, which is expected since marriage probability appears in the eligibles equation, i.e., (3). Similarly, the age-profile of competitors follows the same pattern since individuals generally compete for partners around their age (i.e., have relatively high $P^s_{x,x'}$ values for a close to a') and the marriage probabilities of these eligibles appear in the competitors equation, i.e., (4). Further, the longer right tails are due to single populations dwindling as age increases (since people are less likely to be never married as they get older).

Second, competition follows the opposite pattern: a U-shape that is right-skewed. Younger individuals face higher competition since their eligibles are disproportionately older than they are and have higher marriage probabilities (see equation (5)). Moreover, younger men face especially high competition because their female eligibles disproportionately marry older men. After their mid-20's, individuals face increasing competition as they age. Indeed, for higher ages, as single populations dwindle, more and more of an individual's eligibles become younger than them. This individual is then competing (on average) with a population that is younger than them, which is relatively larger. For example, for a 40-year-old single man, there are more 35-year-old single women than 40-year-old single women (since 40-year-old women are more likely to be married). Each of these 35-year-old women, in turn, has many more male eligibles than the 40-year-old man has female eligibles.

Third, for a similar reason, men face more competitors across the age distribution. Even as eligibles approximately equalize for men and women around age 31, competitors for men remain well above those for women. Compared to women, men are competing for a relatively younger population, for which there are more individuals. In turn, these eligibles for men have more eligibles of their own, driving up the number of competitors a man faces.

Fourth, the overall competition measure average is not 1, but instead, it is 1.27 (1.54 for men and 1.01 for women, for all ages 18-45).[3] This is not unexpected, as the measure is not designed to enforce symmetry. One should think of the competition measure as a sex-age-location-time-specific tightness in the marriage market, not as a relative success probability centered around 1. If individuals of different ages chase similar partner-types on average, this will push the average competition measure up.[4]

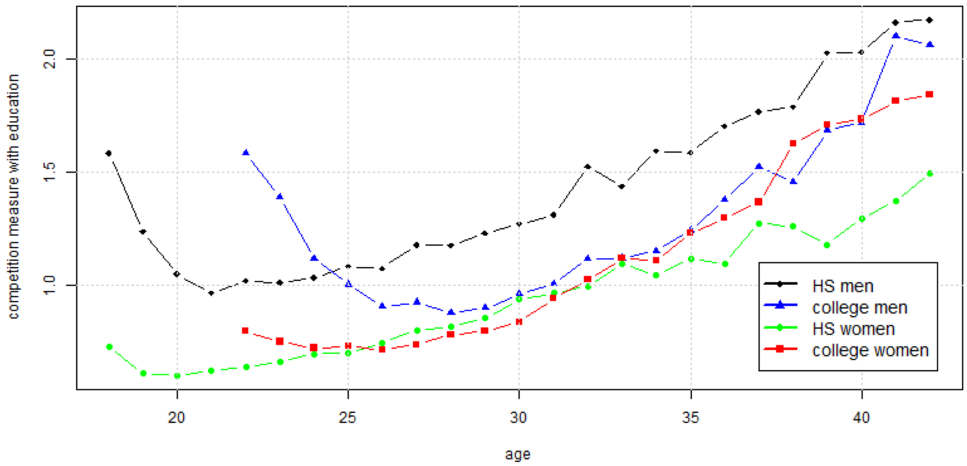
Next, consider Figure A.2, which is analogous to Figure A.1 but with education in the competition measure. Further, we consider the measure broken down by education in addition to sex. Similar to the baseline measure, the age profile for competition is a U-shape with a right skew, and that for competitors and eligibles is an upside-down-U shape with a right skew. Also similar to the baseline measure, men face more competition than women.

The breakdown by education introduces interesting dynamics. At young ages, college men face more competition than high school men. However, as the marriage probability of college men rises considerably above that of high school men, their eligibles increase, and their competition measure falls well below that of high school men. This is true despite college men facing more competitors than high school men. A different pattern emerges for women. While college women start with a higher competition measure, they converge around age 27, and stay approximately the same until college women are slightly larger in the late 30's. Competition and eligibles are greater for college women after age 25.

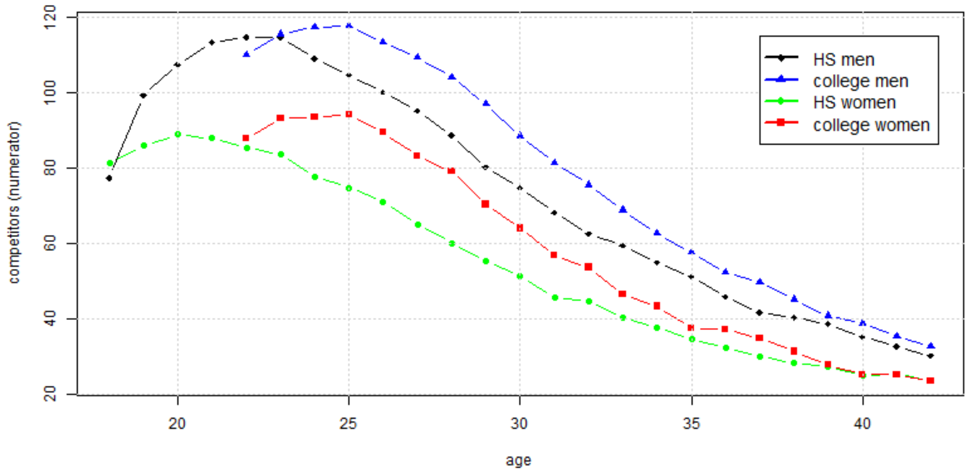
Following Goldsmith-Pinkham et al. (2020), Tables A.2 and A.3 report cross-sectional regressions from the base year (2005) of competition on covariates. For competition measure

[3]The overall average is slightly skewed toward men because there are more never-married men than never-married women. Also, note that these statistics are not reported in a table elsewhere.

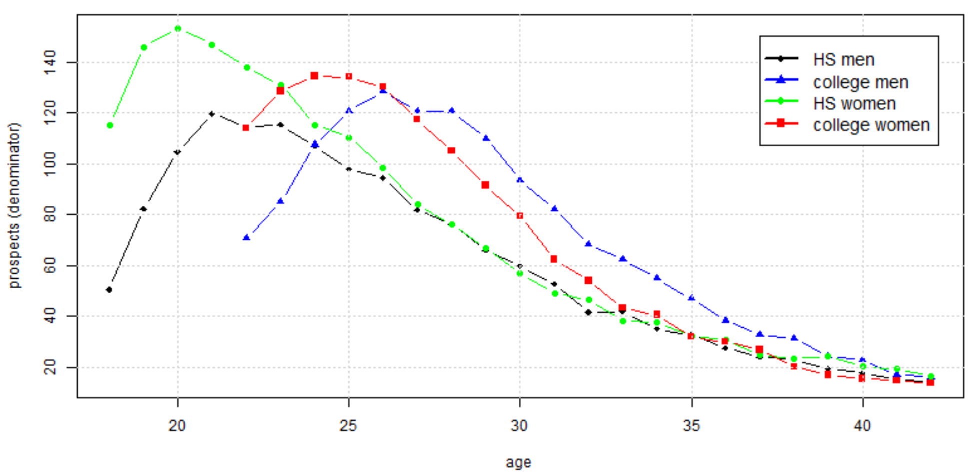
[4]As an extreme example, assume that men of all ages almost exclusively married women of age 30. Each man would have relatively few eligibles (since there are only so many 30-year-old women) and relatively many competitors (since all the other men want to marry 30-year-old women too). Furthermore, every woman who was not 30 would also have relatively few eligibles (since few men want to marry them, their marriage probability would be low), driving their competition measure up. Thus, on average, competition measures would be high for both sexes.



(a) Competition measure with education



(b) Competitors (numerator)



(c) Eligibles (denominator)

Figure A.2: Comp. measure with education, competitors, & eligibles by ed., sex, & age

Notes: Values for 18-21-year-old college individuals are omitted because of sparsity.

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Base year variable								
non-participation rate	0.0135	0.0612^{***}	0.0282^{**}	0.0317^{**}	0.0180^{*}	0.0297^{***}	0.0551^{***}	0.0337^{**}
	(0.0086)	(0.0097)	(0.0141)	(0.0160)	(0.0104)	(0.0101)	(0.0173)	(0.0170)
Fit statistics								
Observations	12,188	10,381	7,125	7,295	12,188	10,381	7,125	7,295
R^2	0.0003	0.0063	0.0008	0.0007	0.0004	0.0013	0.0017	0.0007

(a) regressor: base year value of non-participation rate

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Base year variable								
log hours	-0.0052	0.0072	-0.0373^{**}	0.0546^{**}	0.0085	0.0445^{***}	-0.1430^{***}	0.0381^{*}
	(0.0115)	(0.0140)	(0.0149)	(0.0216)	(0.0135)	(0.0146)	(0.0230)	(0.0219)
Fit statistics								
Observations	10,391	8,384	6,474	6,807	10,391	8,384	6,474	6,807
R^2	2.7×10^{-5}	4.62×10^{-5}	0.0015	0.0018	4.92×10^{-5}	0.0015	0.0120	0.0009

(b) regressor: base year value of log hours

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Base year variable								
log income	0.0343^{***}	0.0220^{***}	0.0553^{***}	0.0664^{***}	0.0682^{***}	0.0644^{***}	-0.0597^{***}	0.0561^{***}
	(0.0039)	(0.0049)	(0.0050)	(0.0064)	(0.0048)	(0.0057)	(0.0073)	(0.0065)
Fit statistics								
Observations	10,391	8,384	6,474	6,807	10,391	8,384	6,474	6,807
R^2	0.0106	0.0034	0.0342	0.0280	0.0282	0.0252	0.0211	0.0195

(c) regressor: base year value of log income

competition measure	**baseline**		**with education**	
subset	*men*	*women*	*men*	*women*
Model:	(1)	(2)	(3)	(4)
Base year variable				
idleness rate	-0.2890^{***}	-0.2154^{***}	-0.2489^{***}	-0.2711^{***}
	(0.0686)	(0.0360)	(0.0626)	(0.0334)
Fit statistics				
Observations	7,486	7,442	7,486	7,442
R^2	0.0027	0.0051	0.0024	0.0098

(d) regressor: base year value of idleness rate

Table A.2: Regressing initial competition measure on covariates (outcome variables)

Notes: For each competition-measure-education-sex combination and for each outcome, we regress the initial competition measure in 2005 on the base year value of the outcome. This sheds light on how high competition locations compare to low competition locations in the base year for different populations (i.e., education-sex combinations). The populations follow those in the main regressions (in Tables 2 and 3). Significance codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Base year variable								
marriage rate	0.1384***	0.0004	0.0999***	0.1151***	0.2384***	-0.0402**	-0.0658***	0.0365***
	(0.0096)	(0.0157)	(0.0098)	(0.0130)	(0.0110)	(0.0159)	(0.0136)	(0.0139)
Fit statistics								
Observations	12,188	10,381	7,125	7,295	12,188	10,381	7,125	7,295
R^2	0.0312	1.84×10^{-7}	0.0237	0.0199	0.0623	0.0014	0.0055	0.0019

(a) regressor: base year value of marriage rate

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Base year variable								
share with bach. deg.	0.2285***	-0.2020***	0.0596**	-0.0101	0.6449***	0.8274***	-0.7716***	0.0686***
	(0.0169)	(0.0259)	(0.0241)	(0.0241)	(0.0204)	(0.0274)	(0.0221)	(0.0252)
Fit statistics								
Observations	12,188	10,381	7,125	7,295	12,188	10,381	7,125	7,295
R^2	0.0227	0.0111	0.0033	7.42×10^{-5}	0.1219	0.1646	0.2946	0.0033

(b) regressor: base year value of share with a bachelor's degree

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Base year variable								
share in metro area	0.0247***	0.0065	0.0332***	-0.0076	0.0677***	0.1110***	-0.1651***	-0.0040
	(0.0052)	(0.0075)	(0.0075)	(0.0092)	(0.0065)	(0.0077)	(0.0109)	(0.0098)
Fit statistics								
Observations	12,188	10,381	7,125	7,295	12,188	10,381	7,125	7,295
R^2	0.0027	0.0001	0.0026	9.02×10^{-5}	0.0139	0.0288	0.0348	2.43×10^{-5}

(c) regressor: base year value of share of population in a metro area

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Base year variable								
non_metro	-0.0267***	-0.0047	-0.0356***	0.0122	-0.0746***	-0.1206***	0.2037***	0.0007
	(0.0062)	(0.0090)	(0.0095)	(0.0112)	(0.0078)	(0.0091)	(0.0133)	(0.0129)
Fit statistics								
Observations	12,188	10,381	7,125	7,295	12,188	10,381	7,125	7,295
R^2	0.0022	4.08×10^{-5}	0.0018	0.0002	0.0118	0.0235	0.0320	5.28×10^{-7}

(d) regressor: base year value of share of population in a non-metro area

Table A.3: Regressing initial competition measure on covariates (non-outcome variables)

Notes: This table is the same as Table A.2, but for non-outcome variables. The share of the population with a bachelor's degree is necessarily at the age-sex-year-location level (as opposed to the education-age-sex-year-location level). The difference between the metro and non-metro panels is meaningful because there is a large third category where the metro status is indeterminable. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

$C^s_{x,\ell,t}$ and covariate $z^s_{x,\ell,t}$, the cross-sectional regression is simply

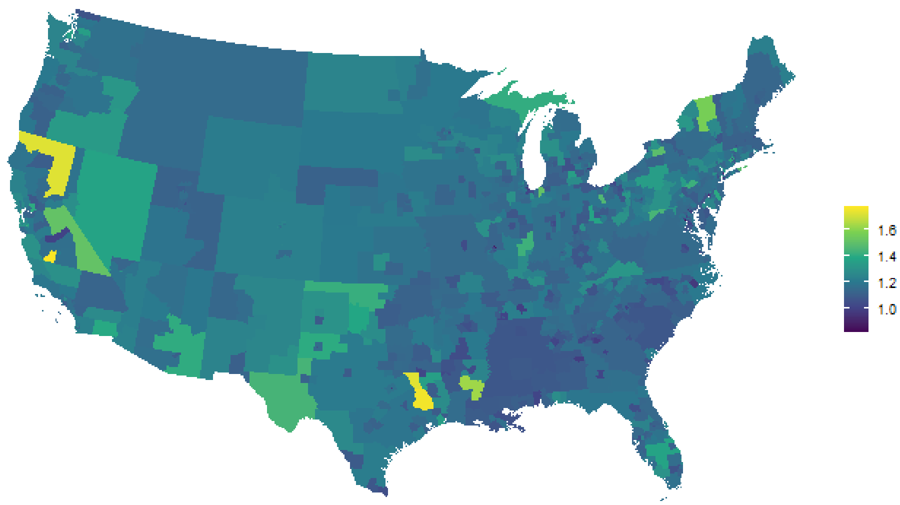
$$\log(C^s_{x,\ell,t_0}) \;=\; \alpha_0 \;+\; \alpha z^s_{x,\ell,t_0} \;+\; \eta_{x,\ell,t}.$$

This exercise is descriptive and intended to document which types of places have high initial competition. For most variables, we offer this calculation for each combination of competition measure (baseline and with education), education level (high school and college), and sex.

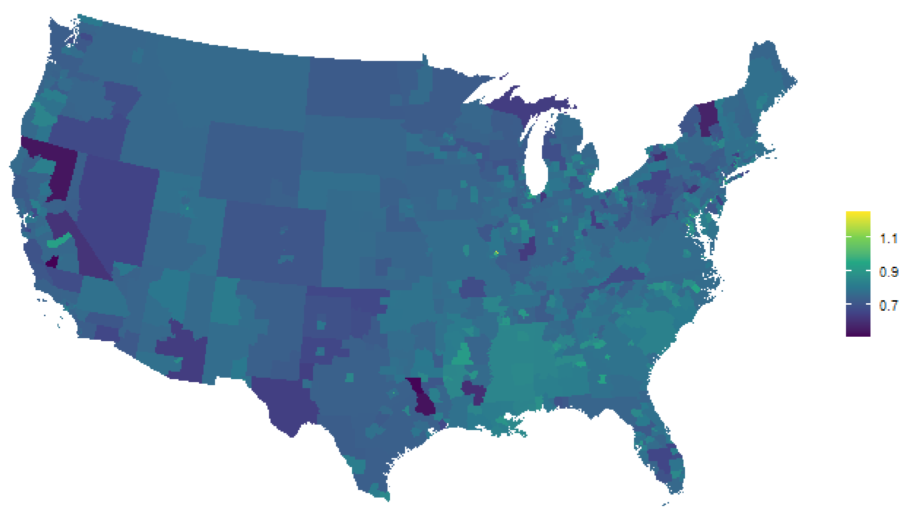
The estimates reveal that areas with initially high competition have slightly higher non-participation rates for all education-sex groups but high school men, lower hours for college men, and higher income for all adults (25-39-year-old and single). One exception is that the college male income estimate flips parity for the measure with education. Idleness is lower across the board (for 18-24-year-old single adults). High school men and college women in these areas have higher marriage rates. Also, for high school men, high-competition areas are more likely to have higher shares of college graduates and are less likely to be metro areas.

Although initially high- and low-competition areas differ on some dimensions, our empirical strategy relies on within-area changes over time, absorbing these baseline differences via fixed effects.

Figure A.3 offers a heat map of the baseline competition measure for men and women separately, where the measure at each location is the population-weighted average over ages and years. For men, there is no strong pattern geographically, although the western U.S. appears to experience slightly more competition on average. For women, the high- and low-competition areas appear to be inverted relative to the men, which is unsurprising given the construction of the competition measure. Additionally, Figure A.4 gives the change in competition over our analysis period 2006-2021.



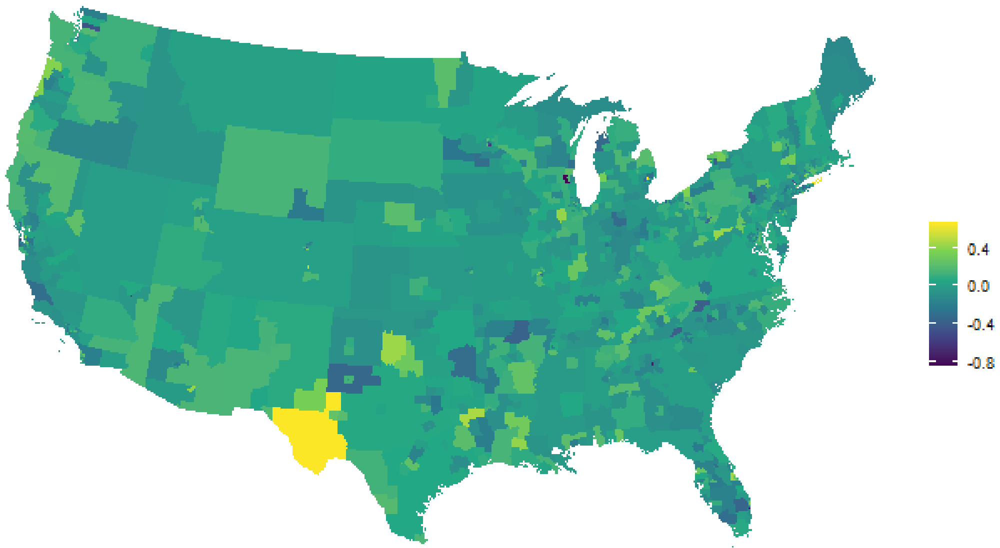
(a) single men 25-39



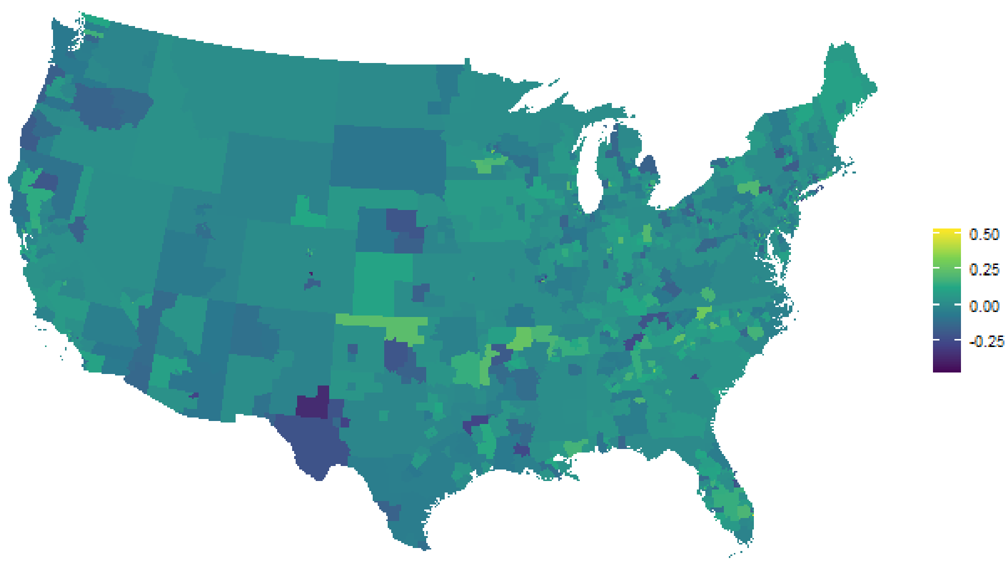
(b) single women 25-39

Figure A.3: Heat map of baseline competition measure

Notes: First, average competition measures at the sex-age-location-year level are taken from all observations of age 25-39-year-olds who report not being in school at the time. Then, one value is assigned to each PUMA and sex by taking the population-weighted average over all ages and years considered.



(a) single men 25-39



(b) single women 25-39

Figure A.4: Heat map of change in baseline competition measure (2021 value minus 2006 value)

Notes: Positive values are places where competition increased more: higher 2021 values than 2006 values.

B Robustness and Extensions

Next, we turn to a robustness analysis of our main results. This analysis is summarized in Section 4 of the body. We also discuss a few extensions.

Note that Appendix D.2 repeats the main analysis for the measure with education, reporting estimates alongside the baseline competition measure. In the robustness exercises and extensions that follow, we continue to report results from both measures.

B.1 Robustness

Differing fixed effects: To evaluate the role of fixed effects, we consider different combinations and measure the effect on the non-participation rate estimates for men from Table 3a. For both male specifications in Table 3a – HS and college men – Table D.8a in Appendix D offers four different fixed-effect combinations: no fixed effects; age, location, and year; age-location and year; age-location and location-year. While this is for the baseline competition measure, Table D.8b does the same for the measure with education. Note that the fourth combination is what we use in the main specification, so these estimates line up with those from Table D.5a.

For all four specifications, the third and fourth combinations (age-location and year, and age-location and location-year, respectively) yield similar 2SLS estimates. However, the first and second combinations (none and age, location, and year, respectively) yield estimates that often differ substantially from the latter two combinations. Further, estimates from these first two combinations often differ from each other. Across all 16 estimates, the instrument is quite strong.

The move from the first combination to the second shows the importance of accounting for consistencies in non-participation across age, location, and time. The move from the second to the third shows the importance of replacing the separate age and location fixed effects with age-location fixed effects. The move from the third to the fourth (i.e., replacing time fixed effects with location-time fixed effects) does not change estimates much.

To gauge whether the added fixed effects remove too much variation, we plot the residuals from 6 of the 16 models from Table D.8 in Figure D.1. In particular, the second, third, and fourth fixed effect combinations for high school men, both for the baseline competition measure and the measure with education. As we increase fixed effects from (2) to (3) to (4), residual variation decreases as expected, i.e., the density becomes more concentrated around zero. However, significant variation remains in the richest model (Figures D.1e and D.1f) – with age-location and location-year fixed effects, which is the combination used in the main results.

Pre-existing trends: It is possible that cells (at the sex-age-location level or education-sex-age-location level) with high and low initial competition levels were on different trends before our analysis period. In that case, it is possible that our main estimates are picking up these trends. We first gauge this possibility with the following simple regression:

$$\upsilon^s_{x,\ell,t} \;=\; \alpha_0 \;+\; \alpha C_{x,\ell,t_0} \;+\; \eta_{x,\ell,t},$$

where $\upsilon^s_{x,\ell,t_1} = y_{x,\ell,t_1} - y_{x,\ell,t_0}$ is a long difference, t_1 is 2019, and t_0 is 2005. 2019 is chosen to omit the effects of COVID-19 on the long difference. The results are in Table D.9.

There are three sets of significant results worth mentioning. First, the high school female non-participation estimates are significantly positive. This suggests that cells where high school women initially faced high marriage competition were already experiencing a positive trend in non-participation (relative to cells initially facing low competition). Second, hours and income estimates for all education-sex combinations except college men are positive, suggesting a preexisting positive trend for these groups in high-competition areas.[5] Magnitudes are quite large. Finally, idleness estimates for men are significantly negative, although magnitudes are relatively small.

[5]Curiously, the college-male estimates flip from significantly negative for the baseline competition measure to significantly positive for the competition measure with education for both hours and income.

To further investigate initial trends, we use a stronger specification to allow for flexible time paths. Considering all years in our analysis period (2006-2021, not just 2019), we interact the initial competition measure with time:

$$y^s_{x,\ell,t} \;=\; \alpha_0 \;+\; \alpha_\tau(t \times C_{x,\ell,t_0}) \;+\; \eta_{x,\ell,t}.$$

The results of this specification are in Table D.10. For several of our main results, the corresponding estimates of α_τ are statistically significant and in the same direction, which suggests the corresponding results from Tables 2 and 3 may reflect long-term trends. For the increase in high school female non-participation and college male hours, the estimates are positive and statistically significant, but only for the measure with education. For the college male hours increase, the estimate for the baseline measure is significantly positive, but that for the measure with education is not (at the 5% level). Lastly, for the male idleness decrease, the estimate for both measures is significantly negative.

While Tables D.9 and D.10 show that initial competition predicts subsequent changes for certain outcomes, they only rely on cross-sectional variation in the base-year competition measure. In contrast, our main specifications identify effects from within-location changes in competition over time while absorbing age-location and location-time fixed effects. To further gauge the role of pre-existing trends in our main estimates, we add a time trend to the age-location fixed effects in our main regression. In particular, the 2SLS regression in equations (8) and (9) becomes

$$\begin{aligned} y^s_{x,\ell,t} &= \beta_0 + \beta \;\log(\hat{C}^s_{x,\ell,t}) + t \times \tau_{\text{age},\ell} + \lambda_{\text{age},\ell} + \tau_{\ell,t} + \epsilon_{x,\ell,t}, \\ \log(C^s_{x,\ell,t}) &= \gamma_0 + \gamma \;\log(\mathcal{C}^s_{x,\ell,t}) + t \times \tau_{\text{age},\ell} + \lambda_{\text{age},\ell} + \tau_{\ell,t} + \nu_{x,\ell,t}. \end{aligned}$$

The results are in Table D.11.

The significantly positive result for female non-participation and the significantly negative result for male idleness remain (even with larger standard errors). This suggests that the relevant estimates from Table 3 are not fully driven by differential age–location trends. Further, it provides strong evidence that the main estimates reflect competition effects as opposed to just slow underlying trends. However, the positive results disappear for college male hours and income, and actually become negative. The significance of the high school male hours also disappears.

This disappearance of positive results does not necessarily nullify the main estimates. Without age–location trends, identification comes from long-run demographic shifts – i.e., gradual changes in the local population composition – and short-run fluctuations – i.e., year-to-year changes in local demographic composition. When we add trends, gradual demographic evolution within each age–location cell is absorbed, and the remaining identifying variation comes only from the short-run fluctuations around those trends. Because most of the variation generated by your Bartik-style demographic instrument occurs through slow-moving population changes, adding trends removes much of the variation identifying the effect. This makes estimates less precise, which can result in a loss of statistical significance, even if the underlying relationship is still present. This is what we see, as standard errors increase sizably.

Alternative geographical units: In our main analysis, we use "consistent PUMAs" as the geographical unit since that is the most granular geography offered in the ACS that is consistent over the time period. However, a less granular geography may be more appropriate. For example, Los Angeles, California is made up of many different PUMAs. Consider two PUMAs on opposite sides of Los Angeles. It is reasonable to think of individuals in those two counties as being in the same "marriage market."

For this reason, we consider two alternate ways to divide locations that are generally less granular than PUMAs. First, we use 2020 definitions of commuting zones (CZs) from the U.S. Department of Agriculture's Economic Research Service (USDA ERS 2026). Second, we use "spillover-defined local labor markets" (SLMs) from Bartik (2024). Both CZs and SLMs are strictly coarser than counties. However, PUMAs are sometimes contained within counties; other times, they consist of multiple counties.

To proceed, we need a way to map the ACS geographical unit to counties. For observations from 2005-2011, we assign each PUMA (not the consistent PUMAs we use in the main analysis, but the PUMAs from the 2000 Census) county weights corresponding to the fraction of the 2010 population in each county around the country. For observations from 2012-2021, we do the same, but use PUMAs from the 2010 Census. This correspondence is given by the "Geocorr 2018: Geographic Correspondence Engine" from the Missouri Census Data Center (MCDC 2018). Then, we map each PUMA to the county with the highest weight. From there, since both CZs and SLMs are strictly coarser than counties, we can then map each PUMA to a unique CZ and SLM.

Results from these alternative geographies are in Tables D.12 and D.13, respectively. The results are very close to those from Section 3. In particular, specifications with significant estimates in Section 3 also have significant estimates here.

Leave-one-out (LOO): Recall that our instrument uses an approximation of local populations using national-level growth. A common concern for Bartik-style instruments like this is that the location's population can mechanically influence national figures (since the location makes up part of the nation). This is potentially problematic since we want the instrument to reflect exogenous national trends as opposed to local shocks.

Since we use many locations (1,078 PUMAs) mostly around the same size, such mechanical endogeneity is less of a concern in our case. However, to be safe, we recompute the results using a leave-one-out (LOO) Bartik construction.[6] The recomputed estimates are in Table D.14. The results are again very close to those from Section 3, which gives us confidence

[6] The true national-level growth used in Section 2.3, $g^m_{x,t}$ for men, is set according to the following equation:

$$\sum_{\ell} M_{x,\ell,t} \;=\; (1 + g^m_{x,t}) \sum_{\ell} M_{x,\ell,t_0}.$$

The LOO construction for location ℓ' uses $\gamma^m_{x,\ell,t}$, which is set according to:

$$\sum_{\ell \neq \ell'} M_{x,\ell,t} \;=\; (1 + \gamma^m_{x,\ell,t}) \sum_{\ell \neq \ell'} M_{x,\ell,t_0}.$$

that mechanical endogeneity plays a negligible role.

Courtship: The goal of the competition measure is to capture how easy it is for men (women) of a particular age in a particular place to find a woman (man) to marry, relative to other places with different population distributions. A potential problem with the measure as presented is that it is not forward-looking. For an age-a single man (woman), the measure uses the age distribution of marriages for an age-a man (woman). In reality, if the person is actually "single" – never married and no boyfriend or girlfriend – and looking for a marriage partner, they are likely not thinking about getting married in the current year (at least on average). They are thinking about marriage a few years in the future.

To account for this, we adjust the competition measure by replacing $\mathcal{P}^s_{x,x'}$ and $P^s_{x,x'}$ with $\mathcal{P}^s_{a+c,a'+c}$ and $P^s_{a+c,a'+c}$, respectively, where c stands for "courtship."[7] The idea is that if an individual meets a potential marriage partner this year, they expect it to take c years before marriage. In practice, we take c to be 3.[8] The results from this exercise are in Table D.15. They are generally consistent with the main results in Section 3, although standard errors are generally slightly bigger.

Long difference: For further robustness, we consider a "long difference" approach. Instead of using the log competition value as the endogenous regressor (and the log of the approximated competition value as the instrument) in the 2SLS setup, we use the change in competition from a base year (and the change in the approximated competition). The

[7]The new competition measure for men is

$$C^m_{a,\ell,t} \;=\; \sum_{a'} P^m_{a+c,a'+c} \sum_{a''} \mathcal{P}^w_{a'+c,a''+c}\, M_{a'',\ell,t} \Big/ \sum_{a'} \mathcal{P}^m_{a+c,a'+c}\, W_{a',\ell,t},$$

and the new approximation is

$$\mathcal{C}^m_{a,\ell,t} \;=\; \sum_{a'} \mathcal{P}^m_{a+c,a'+c} \sum_{a''} P^w_{a'+c,a''+c}(1+g^m_{a'',t}) M_{a'',\ell,t_0} \Big/ \sum_{a'} \mathcal{P}^m_{a+c,a'+c}(1+g^w_{a',t}) W_{a',\ell,t_0}.$$

Those for women are analogously defined. Note that the individual is considering the stock of men and women from the current year.

[8]Note two caveats. First, we could not find data to inform the value of c in a principled way. Second, in the data, we cannot see whether someone who is "single" (never married and not cohabitating with an unmarried partner) is already dating someone seriously or engaged. This is a problem whether or not we adjust for courtship. However, as mentioned, we do remove cohabitants from the single population.

outcome variable also becomes a difference. In particular, take $\upsilon^s_{x,\ell,t}$ to be the change in non-participation rate from a base year t_0, i.e., the "long difference" in the non-participation rate. So if $y^s_{x,\ell,t}$ is the non-participation rate, then $\upsilon^s_{x,\ell,t} = y^s_{x,\ell,t} - y^s_{x,\ell,t_0}$. Also, take $\Delta^s_{x,\ell,t}$ to be the long difference in the competition measure ($\Delta^s_{x,\ell,t} = C^s_{x,\ell,t} - C^s_{x,\ell,t_0}$) and $\delta^s_{x,\ell,t}$ to be the long difference in the approximation ($\delta^s_{x,\ell,t} = \mathcal{C}^s_{x,\ell,t} - \mathcal{C}^s_{x,\ell,t_0}$).

With this notation in place, consider the following two-stage least squares (2SLS) regression:

$$\begin{aligned}
\upsilon^s_{x,\ell,t} &= \beta_0 + \beta\ \Delta^s_{x,\ell,t} + \lambda_{\text{age},\ell} + \tau_{\ell,t} + \epsilon_{x,\ell,t}, \\
\Delta^s_{x,\ell,t} &= \gamma_0 + \gamma\ \delta^s_{x,\ell,t}\ + \lambda_{\text{age},\ell} + \tau_{\ell,t} + \nu_{x,\ell,t}.
\end{aligned}$$

β (if interpreted causally) would be the effect of a 0.01 increase in the competition measure on the change in the non-participation rate.

The results are reported in Table D.16. Patterns in estimates match the main analysis, and coefficients are generally similar despite a slightly different interpretation.[9]

B.2 Summarizing robust estimates

In the main analysis, the following effects were significant at the 5% level: for high school non-participation, a decrease for men and an increase for females; for hours, a decrease for high school men and an increase for college men; for income, an increase for college men; and for idleness, a decrease for high school men. Here, we briefly summarize how robust these estimates are to the above specifications. To assess robustness, we consider the time trend analysis, the alternate geographies (CZs and SLMs), the LOO test, the courtship test, and the long difference test using the baseline measure, and the main analysis repeated using the

[9] For example, in Table 3a, the female non-participation 2SLS estimate using the baseline competition measure (0.11) has the following interpretation: a 1% increase in the competition measure decreases non-participation by 0.11 pp on average. In contrast, the corresponding 2SLS coefficient (0.10 pp) from Table D.16a has the following interpretation: a 0.01 increase in the competition measure decreases non-participation by 0.10 pp on average. Since the competition measure has an average greater than 1, 1% is generally bigger than 0.01, so we would expect long-difference coefficients to be slightly attenuated.

measure with education. In particular, we do not consider estimates from the measure with education for the robustness tests (although they are reported alongside estimates from the baseline measure).

For both alternative geographies and the LOO and courtship tests, all effects remain significant at the 5% level. For the long difference test and the measure with education, all effects also remain significant at the 5% level, except the high school male non-participation estimate, which is only significant at the 10% level. The results of the time trend analysis suggest that the non-participation effect for high school women (25-39) and the idleness effect for men (18-24) are the most broadly robust. However, this analysis has significant precision issues. For this reason, we still consider the three effects "passing" the other tests (the hours and income increase for college men, and the hours decrease for high school men) to be fairly robust.

B.3 Extensions

We now consider a few extensions. These specifications are not directly relevant to robustness of the main results because they consider different populations or approaches.

Splitting competition into competitors and eligibles: Recall that the competition measure is a ratio of competitors to eligibles: from equation (4),

$$C^m_{x,\ell,t} \;=\; \text{competitors}^m_{x,\ell,t} \;\big/\; \text{eligibles}^m_{x,\ell,t}.$$

As an extension to the 2SLS specification in (8) and (9), we consider regressions with competitors and eligibles separated:

$$\begin{aligned}
y^s_{x,\ell,t} &= \beta_0 + \beta\ \log(\widehat{\text{competitors}}^s_{x,\ell,t}) + \log(\widehat{\text{eligibles}}^s_{x,\ell,t}) + \lambda_{\text{age},\ell} + \tau_{\ell,t} + \epsilon_{x,\ell,t}, \\
\log(\text{competitors}^s_{x,\ell,t}) &= \gamma_0 + \gamma \log(\overline{\text{competitors}}^s_{x,\ell,t}) + \log(\overline{\text{eligibles}}^s_{x,\ell,t}) + \lambda_{\text{age},\ell} + \tau_{\ell,t} + \nu_{x,\ell,t}. \\
\log(\text{eligibles}^s_{x,\ell,t}) &= \delta_0 + \delta \log(\overline{\text{competitors}}^s_{x,\ell,t}) + \log(\overline{\text{eligibles}}^s_{x,\ell,t}) + \lambda_{\text{age},\ell} + \tau_{\ell,t} + \eta_{x,\ell,t},
\end{aligned}$$

where $\overline{\text{competitors}}^{s}_{x,\ell,t}$ and $\overline{\text{eligibles}}^{s}_{x,\ell,t}$ are the approximations of $\text{competitors}^{s}_{x,\ell,t}$ and $\text{eligibles}^{s}_{x,\ell,t}$ (respectively) defined in the numerator and denominator (respectively) of (7), and the latter two equations are the two first-stage equations. Note that since population shares are used to construct the competitor and eligible terms (and their approximations), the values are comparable across locations with different population sizes.

The results of these regressions are reported in Tables D.17 and D.18. Generally, when there is a significantly positive (negative) coefficient on competition in the main results (Tables 2 and 3), there is a significantly positive (negative) coefficient on competitors and a significantly negative (positive) coefficient on eligibles. This is what we would expect since competition is the ratio of competitors to eligibles. For example, for college male income, the coefficients on competitors and eligibles are significantly positive and negative, respectively, coinciding with the significantly positive estimate on competition. However, this is not the case for college male hours. Further, significant estimates on competitors and eligibles sometimes show up when the estimate on competition is not significant (for example, college female income). For male idleness, there is a significantly positive estimate on eligibles, but not a significantly negative estimate on competitors, suggesting that the negative estimate on competition is likely driven by eligibles rather than competitors.

Splitting up time period: Since 2006-2021 is a relatively long time period – covering the Great Recession and subsequent recovery, the expansion of dating apps, COVID-19, etc. – we split it in half and rerun the analysis on each time period: 2006-2013 and 2014-2021. The results are in Tables D.19 and D.20, respectively.

One immediate drawback of cutting the sample in half is losing statistical power. Comparing to the main results in Table D.5, standard errors increase across the board. While there is variance, the increase is generally around a factor of two.

Some of the previous results hold up across both periods. The non-participation rate estimates for high school women are significantly positive at the 10% level for both time periods (and both competition measures). Also, the idleness estimates for younger men are significantly negative at the 5% level across specifications.

Other results hold up in only one of the periods. The hours estimates for high school men are significantly negative (for both measures) for 2006-2013, but not for 2014-2021. Also, the college male income estimates are significantly positive in the later period, but are insignificant in the earlier period (although the significantly positive estimates are smaller than in the main analysis). Still other result disappear entirely. The college male hours estimates are insignificant across the board.

When there are time differences – like for high school male hours and college male income – this could help clarify *when* the effect observed in the main results occurs. When significance disappears in one or both of the time periods, this could be due in large part to a loss in precision.[10] Also, because 2SLS estimators are nonlinear and depend on the covariance structure between the instrument, endogenous regressor, and outcome variable, the pooled estimates are not mechanically equal to averages of the split-sample estimates.

Finally, the time split shows countervailing effects that were not visible in the aggregate estimates. For high school male income, despite no significance for the overall period, estimates (from both measures) are significantly negative in the earlier period and significantly positive in the later period.

In summary, the split-sample analysis suggests that some effects are stable across periods, while others vary meaningfully over time. Differences across periods may reflect genuine changes in the relationship between marriage-market competition and labor supply across distinct economic and social environments. Further, the significance of some estimates disappears in one or both periods, which appears to be driven by halving the sample.

When a sufficient amount of precision is lost for certain variables, the utility of 2SLS is depleted, and true effects are likely poorly estimated. Thus, lack of significance here should

[10]For example, consider the college male income estimates from 2006-2013. Both 2SLS estimates are large, but standard errors are so large that they are insignificant, and, further, the Wu-Hausman p-values suggest the OLS estimates are preferred. OLS estimates are close to zero, likely because of the previously-mentioned problem of attenuation bias. This lack of precision likely also plays a role in 2014-2021 estimates being smaller than estimates from the whole period (in Table 2).

not necessarily be interpreted as evidence of no effect.

Unmarried: Next, we use unmarried individuals as opposed to "single" individuals (i.e., never married and not cohabitating with an unmarried partner). "Unmarried" individuals include those who have been married before, but are separated, divorced, or widowed, and those who are cohabitating with an unmarried partner. Since this is a different population, we expect some differences in results.

The results are in Table D.21. Estimates are broadly consistent with the main results, but a few things are worth noting. First, estimates for high school female non-participation and male idleness are notably bigger and remain significant. Second, high school male hours estimates are smaller and no longer significant. Third, high school female income estimates are now significantly negative. Fourth, college male income estimates are attenuated, but still significant.

Removing those with children: We now consider "single" individuals without a child, what we call the "childless" specification. The results are in Table D.22. Patterns are broadly the same as the main results, but statistical significance is attenuated for high school male and female non-participation, and high school male hours. Significance holds up for college male income and hours (25-39) and male idleness (18-24).

White population only: Finally, we run a specification that computes all model objects using only white observations. Not only are outcomes and populations computed using only whites, but marriage probability distributions are only computed using newlywed whites who marry other whites. The results are in Table D.23. The male idleness decrease (18-24) and the college male hours and income increases (25-39) remain significant. However, statistical significance for high school male and female nonparticipation and high school male hours attenuates or disappears. This suggests that non-white observations and interracial marriages may be driving those results. We leave a further investigation for future research.

C Possible interpretations of results

Male results: As mentioned, competition may be a motivator for men to become more economically attractive. This mechanism has support in the relevant theoretical and empirical literature. We give some examples here. In Gary Becker's foundation models of the marriage market (Becker 1973; Becker 1974), traits that increase a single man's expected household output (like earnings potential), increase matching value, making them a more desirable (or economically attractive) spouse. Exploiting a major migration episode in the U.S., Angrist (2002) finds that higher sex ratios increase male earnings.[11] Using data from an online dating site, Hitsch et al. (2010) find that, on average, marriage-seeking women place significant importance on income and have a strong preference for an educated partner. Further analysis supports the idea that observable preferences in dating markets can generalize to marriage outcomes.

An alternative explanation is that men may be self-insuring against income volatility, since they could marry a spouse who brings in income. The empirical literature is clear that marriage often functions as an economic risk-sharing institution (Blundell et al. 2016; Pierre-André Chiappori et al. 2018). So if a man perceives he is less likely to get married or will get married later – and, thus, not have someone to share risk with – he may hedge against economic risk by increasing labor supply and income. Also, putting income *volatility* aside, a man may adjust to lower *expected* future household income by earning more himself in the present. In Section 3.2, we directly test whether our measures of marriage competition decrease the probability of marriage. Across all specifications for 25-39-year-olds, competition significantly decreases marriage probability. This lends some support to a self-insurance mechanism.

Furthermore, such a mechanism seems especially plausible among men competing for women likely to work while married. Since within-education-group marriages are more com-

[11] He further finds that higher sex ratios increase the earnings of men with young children. While these results are not specifically for single men, he concludes they may reflect greater investment in education or on-the-job training when the marriage market became more competitive.

mon and college women have high labor force attachment, the college male income and hours increase may reflect these mechanisms. To test this possibility, we add an interaction of the competition measure ($\log(\hat{C}^s_{x,\ell,t})$) with the education-location-time-specific fraction of heterosexual marriages where both spouses are in the labor force ($\omega^s_{ed,\ell,t}$).[12] The idea is that the higher $\omega^s_{ed,\ell,t}$ is, the more likely an individual thinks both they and their spouse will work if they get married.

Table D.24 reports estimates from the interaction term for the three 25-39-year-old outcomes. While estimates are noisy, the college-male hours estimate is significantly positive at the 5% level. So, loosely speaking, the more likely it is for both spouses to be in the labor force, the larger the hours effect of competition. This lends more support to a self-insurance mechanism. However, the college-male income estimate is insignificant, but again, the estimates are noisy.

Female results: This rise in non-participation for high school women is worth comparing to the opposing result for high school men. At first glance, increased non-participation for women may seem counterintuitive: if a woman is less likely to marry because of heightened competition (which we see in Table 1a), one may expect that they insure themselves through stronger labor force attachment. While this is possibly the case for some women, the estimates suggest it is overcome by an opposing mechanism. One possible opposing mechanism supposes that men's and women's labor market effort serves different signaling roles in the marriage market.

For men, labor market participation and employment stability may be positive marriage market signals, as discussed above. Thus, it is plausible that when competition faced by men rises, some non-participating men are more likely to join the labor force, and some participating men are less likely to leave. On the other hand, for women, labor market participation may be a substitute for traits valued by some men in marriage. Thus, when competition

[12] Note the fractions are sex-specific. $\omega^m_{ed,\ell,t}$ considers all heterosexual marriages where the man is 18-45, and weights using the weight of the male in the marriage. $\omega^w_{ed,\ell,t}$ does the analogous thing but for women. This interaction is appropriately instrumented for in the 2SLS specification.

faced by women rises, some women may shift time toward traditional specialization norms; in particular, some participating women may become more likely to leave the labor force, and some non-participating women may become less likely to join. This mechanism is most plausible for less-educated women – the group for which we find the significant increase in non-participation – because they have less labor market attachment on average.

Furthermore, the estimates in Table D.24 suggest that the effect is *larger* when both spouses are more likely to be in the labor force. Perhaps, in marriage markets where married women are more likely to work, some single women angle to differentiate themselves by leaving the labor force. Much of the relevant literature focuses on female participation *after* marriage, not before marriage.[13] Our focus on single women is part of our contribution.

Tangentially related is the work of Bursztyn et al. (2017), who study questionnaire data from newly admitted students to an elite MBA program. The questionnaire surveyed job preferences and personality traits to be used in internship placement decisions. When respondents believed their classmates would not see their responses, single and non-single women gave similar responses. When respondents believed their classmates *would* see their responses, non-single women answered similarly. However, single women reported lower desired yearly compensation, lower willingness to travel, a willingness to work fewer hours, less professional ambition, and less tendency for leadership. This result suggests some single women may try to appear less career-oriented to prospective marriage partners.[14]

[13] Relevant papers include Grossbard-Shechtman and Neuman (1988), Pierre-Andre Chiappori et al. (2002), Amuedo-Dorantes and Grossbard (2007), Rapoport et al. (2011), Macunovich (2012), and Bertrand et al. (2015). For example, using CPS data from 1965-2005, Amuedo-Dorantes and Grossbard (2007) find that sex ratios are inversely related to participation for married women. Note that this is the opposite direction of our finding: a lower sex ratio (i.e., higher marriage market competition faced by women) is associated with greater participation (less non-participation).

[14] However, connecting it to our result may be a stretch since it is from a unique context. Our finding is for high school women, not college women, let alone women pursuing a graduate degree. Also, our findings are on actual labor force attachment, not stated career ambitions.

D Additional Tables and Figures

D.1 First-stage tables

competition measure	**baseline**		**with education**	
subset	*HS*	*college*	*HS*	*college*
	(1)	(2)	(3)	(4)
First stage				
Competition approximation (log)	0.9665***	0.9892***	0.9270***	0.8197***
	(0.0349)	(0.0428)	(0.0347)	(0.0499)
Fit statistics				
Observations	215,378	133,564	215,378	133,564
Adjusted R^2	0.9008	0.8999	0.9402	0.9021
F-test	12,161.4	7,931.8	12,859.5	2,791.9
Wald	765.7	534.5	712.8	270.2

Table D.1: First-stage results for single men 25-39

Notes: As mentioned in footnote 13, these first-stage results are from the sample used for the non-participation rate, the outcome variable for 25-39-year-olds that imposes the fewest data restrictions. Each of the five specifications is described at the beginning of Section 3. Sig. codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	**baseline**		**with education**	
subset	*HS*	*college*	*HS*	*college*
	(1)	(2)	(3)	(4)
First stage				
Competition approximation (log)	0.9946***	1.137***	1.005***	1.075***
	(0.0290)	(0.0336)	(0.0376)	(0.0305)
Fit statistics				
Observations	183,229	135,788	183,229	135,788
Adjusted R^2	0.9454	0.9391	0.9569	0.9073
F-test	19,491.5	15,380.9	11,115.5	18,177.4
Wald	1,174.9	1,145.5	714.9	1,240.4

Table D.2: First-stage results for single women 25-39

Notes: This table is analogous to Table D.1 but for women. Sig. codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	**baseline**		**with education**	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
First stage				
Competition approximation (log)	0.9810***	0.9972***	0.9424***	0.9697***
	(0.0372)	(0.0326)	(0.0345)	(0.0484)
Fit statistics				
Observations	119,504	117,652	119,504	117,652
Adjusted R^2	0.9802	0.9781	0.9789	0.9721
F-test	4,475.3	8,167.8	4,044.7	2,822.6
Wald	696.1	932.7	748.2	402.1

Table D.3: First-stage results for single men and women 18-24

Notes: As mentioned in footnote 13, these first-stage results are from the idleness rate. Sig. codes: ***: 0.01, **: 0.05, *: 0.1.

D.2 Results for both competition measures

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	-0.0157***	-0.0123***	-0.0454***	-0.0427***	-0.0179***	-0.0130***	-0.0327***	-0.0374***
	(0.0034)	(0.0035)	(0.0055)	(0.0057)	(0.0036)	(0.0035)	(0.0037)	(0.0048)
2SLS								
Competition (log)	-0.1062***	-0.0744***	-0.3036***	-0.2621***	-0.1156***	-0.0784***	-0.3935***	-0.1374***
	(0.0162)	(0.0122)	(0.0271)	(0.0205)	(0.0179)	(0.0165)	(0.0342)	(0.0163)
Fit statistics								
Observations	219,484	191,932	143,558	147,948	219,484	191,932	143,558	147,948
F-test (1st stage)	12,345.3	19,824.4	8,456.0	15,671.7	13,003.5	11,314.1	2,954.3	18,548.4
Wald (1st stage)	784.1	1,200.2	582.5	1,199.4	721.6	736.1	284.8	1,245.1
Wu-Hausman, p-value	4.72×10^{-11}	6.2×10^{-8}	5.92×10^{-28}	7.95×10^{-36}	6.14×10^{-11}	4.4×10^{-5}	1.07×10^{-39}	1.22×10^{-14}

(a) 25-39-year-old single adults

competition measure	baseline		with education	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	-0.0087***	-0.0148***	-0.0046***	-0.0049**
	(0.0015)	(0.0025)	(0.0013)	(0.0020)
2SLS				
Competition (log)	-0.0286***	-0.1097***	0.0066*	0.0058
	(0.0088)	(0.0120)	(0.0039)	(0.0048)
Fit statistics				
Observations	120,231	119,610	120,231	119,610
F-test (1st stage), Competition (log)	4,555.9	8,532.6	22,142.1	32,086.4
Wald (1st stage), Competition (log)	708.4	971.3	4,180.6	6,934.2
Wu-Hausman, p-value	0.0062	3.28×10^{-23}	0.0001	0.0046

(b) 18-24-year-old single adults

Table D.4: Effects of marriage competition on marriage probability

Notes: This table gives OLS and 2SLS results for the effect of marriage market competition on the probability of getting married in the next year. The top panel is for 25-39-year-old single adults broken down by measure type (baseline or with education), education (high school or college), and sex. The bottom panel is for 18-24-year-old single adults broken down by measure type and sex only. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	-0.0049	0.0238*	0.0060	0.0154*	-0.0096	0.0232	0.0027	0.0122*
	(0.0108)	(0.0135)	(0.0102)	(0.0085)	(0.0118)	(0.0143)	(0.0067)	(0.0067)
2SLS								
Competition (log)	-0.0991**	0.1076**	-0.0803*	0.0178	-0.0956*	0.1552**	-0.0814*	-0.0018
	(0.0455)	(0.0465)	(0.0415)	(0.0273)	(0.0505)	(0.0633)	(0.0482)	(0.0193)
Fit statistics								
Observations	215,378	183,229	133,564	135,788	215,378	183,229	133,564	135,788
F-test (1st stage)	12,161.4	19,491.5	7,931.8	15,380.9	12,859.5	11,115.5	2,791.9	18,177.4
Wald (1st stage)	765.7	1,174.9	534.5	1,145.5	712.8	714.9	270.2	1,240.4
Wu-Hausman, p-value	0.0194	0.0233	0.0131	0.9156	0.0482	0.0105	0.0376	0.4021

(a) Non-participation rate for 25-39-year-old single adults

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0109	0.0135	0.0022	-0.0104	0.0141	0.0185	-0.0003	-0.0088
	(0.0092)	(0.0119)	(0.0107)	(0.0103)	(0.0103)	(0.0127)	(0.0076)	(0.0086)
2SLS								
Competition (log)	-0.1011**	-0.0081	0.1139**	-0.0083	-0.1217***	-0.0115	0.1273**	-0.0233
	(0.0393)	(0.0332)	(0.0458)	(0.0293)	(0.0431)	(0.0466)	(0.0519)	(0.0223)
Fit statistics								
Observations	178,640	149,872	122,626	128,069	178,640	149,872	122,626	128,069
F-test (1st stage)	10,121.1	16,454.4	7,417.7	14,524.9	10,447.1	9,177.3	2,668.0	17,276.9
Wald (1st stage)	687.4	1,076.8	493.9	1,105.4	631.9	631.1	250.4	1,197.1
Wu-Hausman, p-value	0.0010	0.4855	0.0048	0.9397	0.0003	0.4926	0.0051	0.4593

(b) log hours for 25-39-year-old single adults

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0185	0.0369	0.0382	-0.0126	0.0082	0.0222	0.0104	-0.0158
	(0.0239)	(0.0289)	(0.0347)	(0.0300)	(0.0274)	(0.0312)	(0.0231)	(0.0237)
2SLS								
Competition (log)	-0.0297	-0.1027	1.119***	0.1284	-0.0406	-0.1368	1.118***	0.0551
	(0.1189)	(0.1036)	(0.1494)	(0.0919)	(0.1353)	(0.1363)	(0.1684)	(0.0705)
Fit statistics								
Observations	178,640	149,872	122,626	128,069	178,640	149,872	122,626	128,069
F-test (1st stage)	10,121.1	16,454.4	7,417.7	14,524.9	10,447.1	9,177.3	2,668.0	17,276.9
Wald (1st stage)	687.4	1,076.8	493.9	1,105.4	631.9	631.1	250.4	1,197.1
Wu-Hausman, p-value	0.6012	0.0776	4.73×10^{-20}	0.0516	0.6293	0.1542	3.21×10^{-16}	0.1835

(c) log income for 25-39-year-old single adults

competition measure	baseline		with education	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	-0.0015	-0.0077	-0.0251***	-0.0515***
	(0.0041)	(0.0058)	(0.0036)	(0.0044)
2SLS				
Competition (log)	-0.1074***	-0.0165	-0.1486***	-0.1943***
	(0.0237)	(0.0216)	(0.0097)	(0.0097)
Fit statistics				
Observations	120,231	119,610	120,231	119,610
F-test (1st stage)	4,555.9	8,532.6	22,142.1	32,086.4
Wald (1st stage)	708.4	971.3	4,180.6	6,934.2
Wu-Hausman, p-value	1.69×10^{-7}	0.6596	1.59×10^{-52}	1.83×10^{-73}

(d) idleness rate for 18-24-year-old single adults

Table D.5: Effects of marriage competition: baseline measure and measure with education

Notes: This table repeats results from Tables 2 and 3, which are for the baseline competition measure, and gives analogous for the measure with education. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

D.3 Complementing the main analysis

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	-0.0014	0.0215*	0.0043	0.0133*	-0.0044	0.0215*	0.0014	0.0112**
	(0.0098)	(0.0116)	(0.0084)	(0.0069)	(0.0108)	(0.0123)	(0.0055)	(0.0055)
2SLS								
Competition (log)	-0.0846**	0.0926**	-0.0629*	0.0103	-0.0796*	0.1366**	-0.0694*	-0.0043
	(0.0414)	(0.0419)	(0.0344)	(0.0232)	(0.0460)	(0.0569)	(0.0398)	(0.0164)
Fit statistics								
Observations	219,484	191,932	143,558	147,948	219,484	191,932	143,558	147,948
F-test (1st stage)	12,345.3	19,824.4	8,456.0	15,671.7	13,003.5	11,314.1	2,954.3	18,548.4
Wald (1st stage)	784.1	1,200.2	582.5	1,199.4	721.6	736.1	284.8	1,245.1
Wu-Hausman, p-value	0.0249	0.0310	0.0213	0.8719	0.0618	0.0124	0.0365	0.2703

(a) idleness rate for 25-39-year-old single adults

competition measure	baseline		with education	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	0.0029	0.0461***	0.0223***	0.0402***
	(0.0063)	(0.0088)	(0.0059)	(0.0066)
2SLS				
Competition (log)	0.2393***	0.4000***	0.1779***	0.1185***
	(0.0355)	(0.0411)	(0.0153)	(0.0159)
Fit statistics				
Observations	120,231	119,610	120,231	119,610
F-test (1st stage)	4,555.9	8,532.6	22,142.1	32,086.4
Wald (1st stage)	708.4	971.3	4,180.6	6,934.2
Wu-Hausman, p-value	6.77×10^{-16}	1.13×10^{-31}	2.56×10^{-40}	6.08×10^{-11}

(b) college attendance rate for 18-24-year-old single adults

Table D.6: Effects of marriage competition: non-participation and idleness

Notes: This table is analogous to Table 3, but considers the idleness rate for 25-39-year-olds and the college attendance rate (undergraduate or graduate school) for 18-24-year-olds. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	baseline		with education	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	-0.0024	0.0098	0.0088	0.0491***
	(0.0076)	(0.0134)	(0.0066)	(0.0094)
2SLS				
Competition (log)	-0.0720*	0.1137**	0.0281*	0.1176***
	(0.0391)	(0.0544)	(0.0150)	(0.0174)
Fit statistics				
Observations	101,330	92,441	101,330	92,441
F-test (1st stage)	4,198.6	6,139.8	26,392.8	40,595.4
Wald (1st stage)	603.8	846.9	3,860.5	6,951.4
Wu-Hausman, p-value	0.0517	0.0328	0.1163	2.86×10^{-7}

(a) log hours

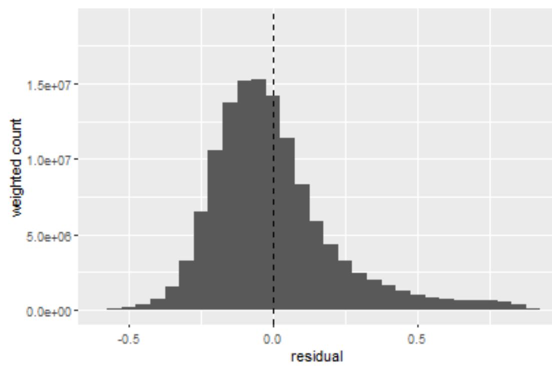
competition measure	baseline		with education	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	-0.0190	0.1175***	0.0936***	0.2664***
	(0.0206)	(0.0329)	(0.0183)	(0.0239)
2SLS				
Competition (log)	-0.2790**	0.6682***	0.3464***	0.6184***
	(0.1140)	(0.1557)	(0.0438)	(0.0459)
Fit statistics				
Observations	101,330	92,441	101,330	92,441
F-test (1st stage)	4,198.6	6,139.8	26,392.8	40,595.4
Wald (1st stage)	603.8	846.9	3,860.5	6,951.4
Wu-Hausman, p-value	0.0050	3.43×10^{-6}	2.19×10^{-15}	2.41×10^{-27}

(b) log income

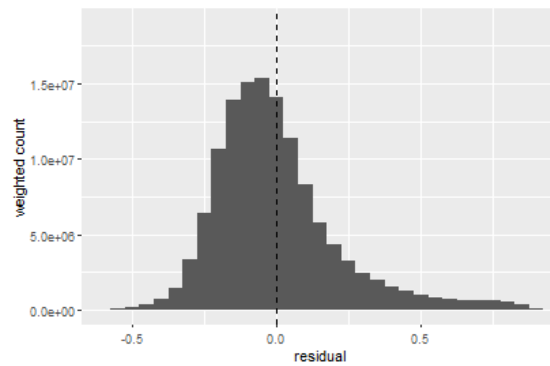
Table D.7: Effects of marriage competition: intensive margin outcomes for 18-24-year-olds

Notes: This table is analogous to Table 2, but for 18-24-year-olds (singles not in school) as opposed to 25-39-year-olds. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

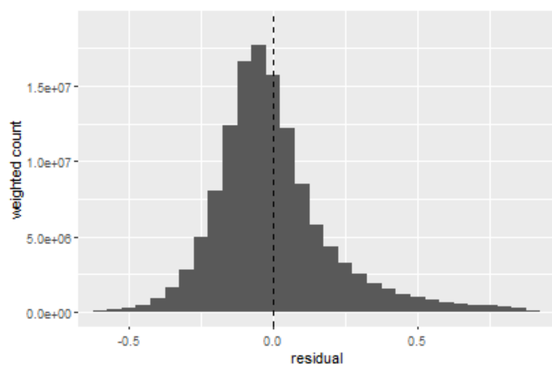
D.4 Alternative fixed effect specifications



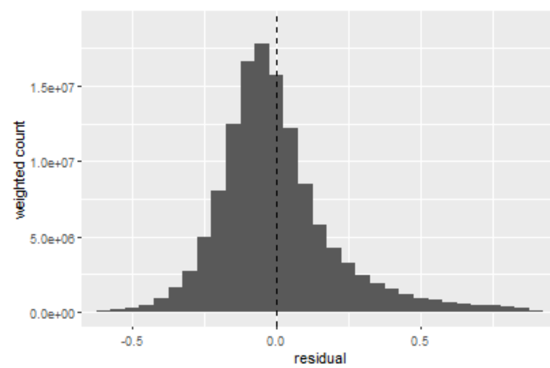
(a) baseline, model (2)



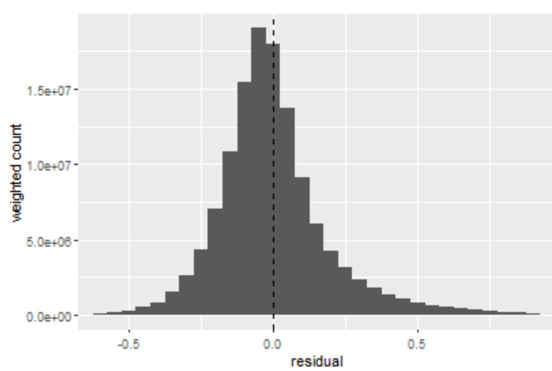
(b) with education, model (2)



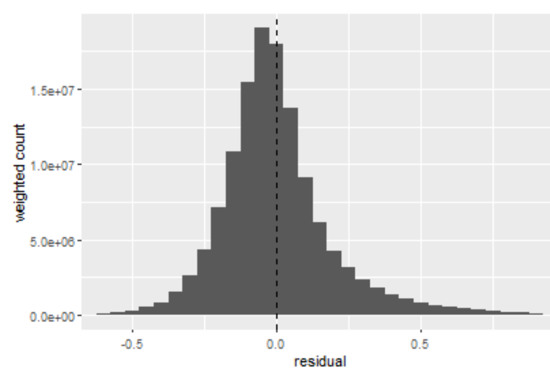
(c) baseline, model (3)



(d) with education, model (3)



(e) baseline, model (4)



(f) with education, model (4)

Figure D.1: Residual histogram plots for HS men non-participation rate

Notes: Consider Table D.8 that differs fixed effects (FE) used in the male non-participation rate. These six plots correspond to the 2SLS regression results from this table for models (2), (3), and (4) for the baseline competition measure and the measure with education. Recall the different FE mixes: model (2) is age, PUMA, and year; model (3) is age-PUMA and year; model (4) is age-PUMA and PUMA-year. Each histogram shows the density of the residual distribution for the corresponding model and competition measure. The plots have the same scale for comparability.

education		HS				college		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First stage								
Comp. approx. (log)	0.5754***	0.1269***	0.8931***	0.9665***	0.5624***	0.1665***	0.9786***	0.9892***
	(0.0141)	(0.0093)	(0.0373)	(0.0349)	(0.0157)	(0.0128)	(0.0464)	(0.0428)
OLS								
Competition (log)	0.1678***	0.0002	-0.0065	-0.0049	0.0425***	0.0057	0.0080	0.0060
	(0.0174)	(0.0070)	(0.0070)	(0.0108)	(0.0039)	(0.0057)	(0.0057)	(0.0102)
2SLS								
Competition (log)	-0.0399*	0.2970***	-0.0887*	-0.0991**	0.0333***	-0.0401	-0.0442	-0.0803*
	(0.0225)	(0.0875)	(0.0477)	(0.0455)	(0.0079)	(0.0494)	(0.0387)	(0.0415)
Fixed-effects								
age		Yes				Yes		
puma		Yes				Yes		
year		Yes	Yes			Yes	Yes	
age-puma			Yes	Yes			Yes	Yes
puma-year				Yes				Yes
Fit statistics								
Observations	215,378	215,378	215,378	215,378	133,564	133,564	133,564	133,564
Adj. R^2 (first stage)	0.3673	0.7080	0.7220	0.9008	0.3558	0.6731	0.6937	0.8999
R^2 (OLS)	0.0127	0.1833	0.2566	0.3360	0.0021	0.0300	0.1784	0.3154
R^2 (2SLS)	-0.0068	0.1716	0.2558	0.3357	0.0020	0.0291	0.1775	0.3147
F-test (1st stage)	125,018.3	1,675.2	3,925.4	12,161.4	73,756.3	1,790.1	2,649.0	7,931.8
Wald (1st stage)	1,669.7	186.6	573.9	765.7	1,287.7	170.1	444.7	534.5
Wu-Hausman p-value	0.0000	0.0000	0.0486	0.0194	0.0066	0.2219	0.1120	0.0131

(a) baseline measure

education		HS				college		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First stage								
Comp. approx. (log)	0.6887***	0.1652***	0.7957***	0.9270***	0.6600***	0.1360***	0.7344***	0.8197***
	(0.0130)	(0.0120)	(0.0384)	(0.0347)	(0.0170)	(0.0141)	(0.0629)	(0.0499)
OLS								
Competition (log)	0.1290***	-0.0087	-0.0115*	-0.0096	0.0363***	-0.0010	0.0037	0.0027
	(0.0135)	(0.0069)	(0.0069)	(0.0118)	(0.0025)	(0.0041)	(0.0042)	(0.0067)
2SLS								
Competition (log)	0.0008	0.2552***	-0.0724	-0.0956*	0.0419***	-0.0487	-0.0263	-0.0814*
	(0.0163)	(0.0751)	(0.0527)	(0.0505)	(0.0038)	(0.0414)	(0.0422)	(0.0482)
Fixed-effects								
age		Yes				Yes		
puma		Yes				Yes		
year		Yes	Yes			Yes	Yes	
age-puma			Yes	Yes			Yes	Yes
puma-year				Yes				Yes
Fit statistics								
Observations	215,378	215,378	215,378	215,378	133,564	133,564	133,564	133,564
Adj. R^2 (first stage)	0.5026	0.7821	0.7959	0.9402	0.4681	0.7384	0.7568	0.9021
R^2 (OLS)	0.0100	0.1833	0.2566	0.3360	0.0035	0.0300	0.1784	0.3154
R^2 (2SLS)	0.0001	0.1741	0.2562	0.3358	0.0034	0.0284	0.1779	0.3140
F-test (1st stage)	217,602.5	2,414.9	3,472.7	12,859.5	117,554.9	1,394.8	1,317.8	2,791.9
Wald (1st stage)	2,812.0	188.5	428.3	712.8	1,513.0	92.51	136.4	270.2
Wu-Hausman p-value	0.0000	0.0000	0.1740	0.0482	0.0017	0.1324	0.3933	0.0376

(b) measure with education

Table D.8: Differing fixed effects: male non-participation rate

Notes: This table considers different combinations of fixed effects for each of the four male specifications in Table 3a. The first four columns in the top panel are for high school men using the baseline measure. The adjacent four columns are for college men. The bottom panel does the same eight columns for the measure with education. Each group of four columns considers four combinations of fixed effects: none; age, location, and year; age-location and year; age-location and location-year. The fourth is our primary specification, so estimates and standard errors from these columns match those in Table D.5a. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

D.5 Pre-existing trends

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Base year variable								
Competition (log)	-0.0371***	0.0421**	-0.0009	0.0239**	-0.0131	0.0449**	0.0147*	0.0188*
	(0.0142)	(0.0187)	(0.0110)	(0.0114)	(0.0111)	(0.0180)	(0.0083)	(0.0114)
Fit statistics								
Observations	13,978	12,103	9,438	9,504	13,978	12,103	9,438	9,504
R^2	0.0005	0.0004	6.86×10^{-7}	0.0005	9.91×10^{-5}	0.0005	0.0003	0.0003

(a) Non-participation rate

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Base year variable								
Competition (log)	0.3013***	0.7875***	-0.2951***	1.445***	0.6043***	1.021***	0.3326***	1.025***
	(0.0692)	(0.0891)	(0.0955)	(0.0995)	(0.0535)	(0.0849)	(0.0720)	(0.1000)
Fit statistics								
Observations	12,026	10,173	8,800	9,064	12,026	10,173	8,800	9,064
R^2	0.0016	0.0076	0.0011	0.0227	0.0105	0.0140	0.0024	0.0114

(b) log hours

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Base year variable								
Competition (log)	0.8321***	2.197***	-0.8446***	4.204***	1.654***	2.861***	0.8327***	3.027***
	(0.1919)	(0.2476)	(0.2732)	(0.2844)	(0.1485)	(0.2358)	(0.2059)	(0.2858)
Fit statistics								
Observations	12,026	10,173	8,800	9,064	12,026	10,173	8,800	9,064
R^2	0.0016	0.0077	0.0011	0.0235	0.0102	0.0143	0.0019	0.0122

(c) income

competition measure	baseline		with education	
subset	*men*	*women*	*men*	*women*
Model:	(1)	(2)	(3)	(4)
Base year variable				
Competition (log)	-0.0059***	-0.0010	-0.0072***	0.0014
	(0.0011)	(0.0039)	(0.0015)	(0.0048)
Fit statistics				
Observations	7,517	7,475	7,517	7,475
R^2	0.0037	8.86×10^{-6}	0.0031	1.08×10^{-5}

(d) idleness rate

Table D.9: Regressing change in outcomes on base year competition measure

Notes: For each competition-measure-education-sex combination and for each outcome, we regress the change in outcome between the base year (2005) and 2019 on the base year competition measure. This helps us understand whether high- and low-competition areas were initially following different trends. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Base year variable								
Competition (log) × year	0.002045**	0.000784	0.000390	-0.000319	0.002329***	0.002972***	0.000293	-0.000104
	(0.000862)	(0.001039)	(0.000641)	(0.000695)	(0.000661)	(0.001082)	(0.000532)	(0.000685)
Fit statistics								
Observations	169,079	129,125	76,724	79,208	169,079	129,125	76,724	79,208
R^2	0.260257	0.139368	0.140776	0.136055	0.260318	0.139459	0.140776	0.136052

(a) Non-participation rate

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Base year variable								
Competition (log) × year	0.001151*	-0.001219	0.001353	-3.1×10^{-5}	0.001025*	-0.001347	0.001833***	-0.000110
	(0.000656)	(0.000827)	(0.000865)	(0.000874)	(0.000523)	(0.000838)	(0.000681)	(0.000860)
Fit statistics								
Observations	126,216	91,526	66,894	71,802	126,216	91,526	66,894	71,802
R^2	0.136175	0.145866	0.205628	0.201032	0.136187	0.145873	0.205741	0.201032

(b) log hours

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Base year variable								
Competition (log) × year	0.003461*	-0.006723***	0.009755***	-0.006551**	0.002669	-0.010139***	0.004851*	-0.004143
	(0.002033)	(0.002542)	(0.003571)	(0.002854)	(0.001622)	(0.002488)	(0.002659)	(0.002744)
Fit statistics								
Observations	126,216	91,526	66,894	71,802	126,216	91,526	66,894	71,802
R^2	0.197424	0.216582	0.411886	0.419850	0.197425	0.216745	0.411750	0.419791

(c) income

competition measure	**baseline**		**with education**	
subset	*men*	*women*	*men*	*women*
Model:	(1)	(2)	(3)	(4)
Base year variable				
Competition (log) × year	-0.000134***	-0.000125	-0.000155***	-3.81×10^{-5}
	(4.32×10^{-5})	(0.000138)	(5.77×10^{-5})	(0.000178)
Fit statistics				
Observations	119,302	118,018	119,302	118,018
R^2	0.256757	0.199292	0.256741	0.199286

(d) idleness rate

Table D.10: Regressing outcomes on base year competition measure interacted with time

Notes: Significance codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	-5.32×10^{-5}	0.0190	0.0081	0.0173*	-0.0031	0.0224	0.0040	0.0146*
	(0.0121)	(0.0154)	(0.0114)	(0.0100)	(0.0134)	(0.0159)	(0.0075)	(0.0080)
2SLS								
Competition (log)	-0.1282	0.2542***	0.0553	0.1153**	-0.1058	0.3638***	0.0060	0.0462
	(0.1006)	(0.0959)	(0.0836)	(0.0525)	(0.1156)	(0.1198)	(0.0678)	(0.0373)
Fit statistics								
Observations	215,378	183,229	133,564	135,788	215,378	183,229	133,564	135,788
F-test (1st stage)	3,127.4	5,671.5	1,992.7	5,942.7	2,956.6	3,846.1	1,359.6	7,021.5
Wald (1st stage)	416.5	506.9	204.2	538.0	290.4	442.7	198.2	440.3
Wu-Hausman, p-value	0.1252	0.0015	0.5198	0.0118	0.2857	0.0003	0.9725	0.2718

(a) Non-participation rate

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0174*	0.0282**	0.0063	-0.0076	0.0200*	0.0311**	0.0032	-0.0081
	(0.0105)	(0.0143)	(0.0134)	(0.0124)	(0.0119)	(0.0153)	(0.0092)	(0.0101)
2SLS								
Competition (log)	0.0254	0.1412*	-0.0347	0.0475	0.0347	0.1380	-0.0590	0.0167
	(0.0798)	(0.0727)	(0.0928)	(0.0536)	(0.0948)	(0.0919)	(0.0766)	(0.0415)
Fit statistics								
Observations	178,640	149,872	122,626	128,069	178,640	149,872	122,626	128,069
F-test (1st stage)	2,707.4	4,699.7	1,826.0	5,571.9	2,385.6	3,120.3	1,305.9	6,666.5
Wald (1st stage)	386.7	456.0	195.2	517.8	260.2	393.4	191.1	419.0
Wu-Hausman, p-value	0.9085	0.0764	0.6315	0.2379	0.8603	0.1895	0.3622	0.4675

(b) log hours

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0391	0.0806**	0.0124	0.0260	0.0203	0.0548	0.0106	0.0165
	(0.0281)	(0.0336)	(0.0415)	(0.0350)	(0.0317)	(0.0365)	(0.0270)	(0.0278)
2SLS								
Competition (log)	0.3281	0.1534	0.1177	0.4557***	0.3779	0.0306	-0.1360	0.4365***
	(0.2145)	(0.2038)	(0.2867)	(0.1644)	(0.2609)	(0.2547)	(0.2290)	(0.1259)
Fit statistics								
Observations	178,640	149,872	122,626	128,069	178,640	149,872	122,626	128,069
F-test (1st stage)	2,707.4	4,699.7	1,826.0	5,571.9	2,385.6	3,120.3	1,305.9	6,666.5
Wald (1st stage)	386.7	456.0	195.2	517.8	260.2	393.4	191.1	419.0
Wu-Hausman, p-value	0.1265	0.6544	0.6796	0.0007	0.1129	0.9069	0.4727	5.82×10^{-6}

(c) income

competition measure	**baseline**		**with education**	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	-0.0018	-0.0042	-0.0258***	-0.0503***
	(0.0044)	(0.0063)	(0.0039)	(0.0047)
2SLS				
Competition (log)	-0.1126***	-0.0286	-0.1427***	-0.1987***
	(0.0310)	(0.0667)	(0.0103)	(0.0106)
Fit statistics				
Observations	120,231	119,610	120,231	119,610
F-test (1st stage)	2,992.9	950.2	21,381.6	29,594.8
Wald (1st stage)	463.0	164.4	4,012.5	6,646.5
Wu-Hausman, p-value	1.6×10^{-5}	0.6994	6.33×10^{-44}	6.41×10^{-70}

(d) idleness rate

Table D.11: Adding a time trend to the age-location fixed effects

Notes: Significance codes: ***: 0.01, **: 0.05, *: 0.1.

D.6 Other robustness tests and extensions

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0064	0.0562**	-0.0108	-0.0002	0.0028	0.0424	-0.0106	-0.0023
	(0.0176)	(0.0233)	(0.0207)	(0.0159)	(0.0203)	(0.0268)	(0.0140)	(0.0131)
2SLS								
Competition (log)	-0.1136**	0.1115**	-0.0790**	0.0137	-0.1180**	0.1655***	-0.0790**	-0.0083
	(0.0471)	(0.0457)	(0.0347)	(0.0275)	(0.0522)	(0.0612)	(0.0395)	(0.0202)
Fit statistics								
Observations	87,726	77,473	53,404	53,249	87,726	77,473	53,404	53,249
F-test (1st stage)	15,263.0	28,641.7	16,432.5	32,359.4	17,421.0	16,852.8	5,399.8	36,237.3
Wald (1st stage)	626.3	513.1	497.3	382.7	552.2	337.0	416.2	637.8
Wu-Hausman, p-value	0.0025	0.0946	0.0375	0.4721	0.0041	0.0111	0.0824	0.6782

(a) Non-participation rate

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	-0.0295**	0.0182	0.0168	-0.0248	-0.0376**	0.0262	-0.0172	-0.0168
	(0.0148)	(0.0178)	(0.0221)	(0.0184)	(0.0168)	(0.0208)	(0.0152)	(0.0148)
2SLS								
Competition (log)	-0.1245***	-0.0174	0.1482***	-0.0409	-0.1444***	-0.0233	0.1407***	-0.0387*
	(0.0394)	(0.0290)	(0.0438)	(0.0277)	(0.0422)	(0.0400)	(0.0481)	(0.0210)
Fit statistics								
Observations	76,019	66,709	49,608	50,791	76,019	66,709	49,608	50,791
F-test (1st stage)	14,731.3	27,108.0	15,743.8	31,154.0	16,557.6	15,556.8	5,211.4	34,891.7
Wald (1st stage)	576.2	514.5	483.4	383.6	512.7	319.1	405.6	633.9
Wu-Hausman, p-value	0.0019	0.1568	0.0002	0.4581	0.0010	0.1841	0.0002	0.1756

(b) log hours

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0371	-0.0023	0.2791***	-0.0087	0.0324	-0.0083	0.0451	-0.0267
	(0.0414)	(0.0596)	(0.0749)	(0.0654)	(0.0469)	(0.0655)	(0.0503)	(0.0530)
2SLS								
Competition (log)	0.0641	-0.1093	1.409***	0.0328	0.0668	-0.1516	1.389***	-0.0079
	(0.1159)	(0.1189)	(0.1199)	(0.1063)	(0.1320)	(0.1540)	(0.1493)	(0.0844)
Fit statistics								
Observations	76,019	66,709	49,608	50,791	76,019	66,709	49,608	50,791
F-test (1st stage)	14,731.3	27,108.0	15,743.8	31,154.0	16,557.6	15,556.8	5,211.4	34,891.7
Wald (1st stage)	576.2	514.5	483.4	383.6	512.7	319.1	405.6	633.9
Wu-Hausman, p-value	0.7462	0.0971	3.15×10^{-27}	0.4884	0.6994	0.1336	1.06×10^{-26}	0.6726

(c) income

competition measure	baseline		with education	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	-0.0091	-0.0040	-0.0407***	-0.0683***
	(0.0076)	(0.0097)	(0.0071)	(0.0081)
2SLS				
Competition (log)	-0.1095***	-0.0243	-0.1538***	-0.2264***
	(0.0230)	(0.0224)	(0.0154)	(0.0145)
Fit statistics				
Observations	44,497	44,262	44,497	44,262
F-test (1st stage)	4,713.3	9,202.0	14,180.4	17,974.0
Wald (1st stage)	411.8	404.5	3,354.7	3,762.6
Wu-Hausman, p-value	1.04×10^{-6}	0.3017	9.92×10^{-25}	1.24×10^{-40}

(d) idleness rate

Table D.12: Alternative geographical unit: commuting zones (CZ)

Notes: Instead of PUMAs, this specification uses CZs as the geographical unit of analysis. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	-0.0061	0.0363**	-0.0076	-0.0062	-0.0079	0.0278	-0.0079	-0.0042
	(0.0132)	(0.0181)	(0.0165)	(0.0146)	(0.0154)	(0.0199)	(0.0107)	(0.0117)
2SLS								
Competition (log)	-0.1052**	0.1153***	-0.0827**	0.0101	-0.0991*	0.1782***	-0.0867**	-0.0124
	(0.0464)	(0.0446)	(0.0339)	(0.0264)	(0.0523)	(0.0614)	(0.0388)	(0.0184)
Fit statistics								
Observations	156,195	127,251	71,156	71,161	156,195	127,251	71,156	71,161
F-test (1st stage)	13,721.2	26,206.9	14,773.4	30,799.9	15,339.8	15,412.1	4,749.4	34,428.8
Wald (1st stage)	637.1	432.0	496.2	359.0	551.2	304.0	405.8	664.3
Wu-Hausman, p-value	0.0130	0.0195	0.0225	0.4005	0.0350	0.0020	0.0419	0.5718

(a) Non-participation rate

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	-0.0137	0.0149	0.0208	-0.0272*	-0.0158	0.0301*	-0.0094	-0.0145
	(0.0121)	(0.0149)	(0.0205)	(0.0163)	(0.0135)	(0.0168)	(0.0123)	(0.0134)
2SLS								
Competition (log)	-0.1075***	-0.0368	0.1471***	-0.0296	-0.1255***	-0.0566	0.1280***	-0.0340*
	(0.0411)	(0.0315)	(0.0445)	(0.0256)	(0.0449)	(0.0427)	(0.0473)	(0.0191)
Fit statistics								
Observations	122,501	100,082	63,944	66,414	122,501	100,082	63,944	66,414
F-test (1st stage)	12,761.4	23,872.7	14,247.7	29,833.8	14,133.3	13,769.9	4,618.9	33,249.9
Wald (1st stage)	565.0	418.7	476.5	361.6	500.2	281.0	396.0	670.7
Wu-Hausman, p-value	0.0031	0.0559	0.0004	0.9181	0.0015	0.0277	0.0011	0.2512

(b) log hours

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0134	-0.0186	0.1839**	0.0300	0.0183	-0.0189	0.0336	0.0070
	(0.0314)	(0.0449)	(0.0743)	(0.0588)	(0.0368)	(0.0488)	(0.0440)	(0.0462)
2SLS								
Competition (log)	0.0415	-0.1468	1.421***	0.0575	0.0628	-0.2281	1.414***	-0.0047
	(0.1166)	(0.1208)	(0.1119)	(0.1001)	(0.1328)	(0.1555)	(0.1289)	(0.0787)
Fit statistics								
Observations	122,501	100,082	63,944	66,414	122,501	100,082	63,944	66,414
F-test (1st stage)	12,761.4	23,872.7	14,247.7	29,833.8	14,133.3	13,769.9	4,618.9	33,249.9
Wald (1st stage)	565.0	418.7	476.5	361.6	500.2	281.0	396.0	670.7
Wu-Hausman, p-value	0.7403	0.0585	9.77×10^{-31}	0.6578	0.6310	0.0339	7.81×10^{-28}	0.8013

(c) income

competition measure	baseline		with education	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	-0.0114**	-0.0027	-0.0391***	-0.0651***
	(0.0053)	(0.0074)	(0.0050)	(0.0057)
2SLS				
Competition (log)	-0.1116***	-0.0279	-0.1522***	-0.2282***
	(0.0226)	(0.0235)	(0.0115)	(0.0113)
Fit statistics				
Observations	85,714	84,845	85,714	84,845
F-test (1st stage)	4,782.1	9,291.8	23,251.1	30,773.0
Wald (1st stage)	474.3	348.6	5,168.6	6,034.1
Wu-Hausman, p-value	1.8×10^{-6}	0.2137	1.5×10^{-38}	1.05×10^{-71}

(d) idleness rate

Table D.13: Alternative geographical unit: spillover-defined local labor market (SLM)

Notes: Instead of PUMAs, this specification uses SLMs as the geographical unit of analysis. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	-0.0049	0.0238*	0.0060	0.0154*	-0.0096	0.0232	0.0027	0.0122*
	(0.0108)	(0.0135)	(0.0102)	(0.0085)	(0.0118)	(0.0143)	(0.0067)	(0.0067)
2SLS								
Competition (log)	-0.1014**	0.1084**	-0.0819*	0.0177	-0.0979*	0.1574**	-0.0858*	-0.0021
	(0.0464)	(0.0469)	(0.0423)	(0.0275)	(0.0514)	(0.0644)	(0.0510)	(0.0195)
Fit statistics								
Observations	215,378	183,229	133,564	135,788	215,378	183,229	133,564	135,788
F-test (1st stage)	11,735.6	19,095.9	7,638.6	15,033.1	12,438.0	10,730.0	2,494.5	17,717.5
Wald (1st stage)	734.2	1,145.0	515.1	1,116.5	686.0	681.6	242.1	1,211.1
Wu-Hausman, p-value	0.0186	0.0234	0.0132	0.9215	0.0461	0.0106	0.0385	0.3986

(a) Non-participation rate

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0109	0.0135	0.0022	-0.0104	0.0141	0.0185	-0.0003	-0.0088
	(0.0092)	(0.0119)	(0.0107)	(0.0103)	(0.0103)	(0.0127)	(0.0076)	(0.0086)
2SLS								
Competition (log)	-0.1029**	-0.0085	0.1165**	-0.0083	-0.1239***	-0.0126	0.1354**	-0.0234
	(0.0400)	(0.0335)	(0.0467)	(0.0296)	(0.0438)	(0.0474)	(0.0549)	(0.0226)
Fit statistics								
Observations	178,640	149,872	122,626	128,069	178,640	149,872	122,626	128,069
F-test (1st stage)	9,763.1	16,117.1	7,142.7	14,192.8	10,101.0	8,854.3	2,381.5	16,834.6
Wald (1st stage)	658.7	1,048.7	476.0	1,076.8	607.6	601.1	224.3	1,167.7
Wu-Hausman, p-value	0.0010	0.4825	0.0046	0.9389	0.0003	0.4851	0.0049	0.4644

(b) log hours

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0185	0.0369	0.0382	-0.0126	0.0082	0.0222	0.0104	-0.0158
	(0.0239)	(0.0289)	(0.0347)	(0.0300)	(0.0274)	(0.0312)	(0.0231)	(0.0237)
2SLS								
Competition (log)	-0.0316	-0.1051	1.142***	0.1298	-0.0422	-0.1420	1.186***	0.0564
	(0.1213)	(0.1046)	(0.1524)	(0.0927)	(0.1379)	(0.1388)	(0.1805)	(0.0712)
Fit statistics								
Observations	178,640	149,872	122,626	128,069	178,640	149,872	122,626	128,069
F-test (1st stage)	9,763.1	16,117.1	7,142.7	14,192.8	10,101.0	8,854.3	2,381.5	16,834.6
Wald (1st stage)	658.7	1,048.7	476.0	1,076.8	607.6	601.1	224.3	1,167.7
Wu-Hausman, p-value	0.5937	0.0757	4.15×10^{-20}	0.0518	0.6239	0.1484	2.75×10^{-16}	0.1815

(c) income

competition measure	baseline		with education	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	-0.0015	-0.0077	-0.0251***	-0.0515***
	(0.0041)	(0.0058)	(0.0036)	(0.0044)
2SLS				
Competition (log)	-0.1101***	-0.0165	-0.1493***	-0.1949***
	(0.0243)	(0.0219)	(0.0098)	(0.0098)
Fit statistics				
Observations	120,231	119,610	120,231	119,610
F-test (1st stage)	4,334.7	8,291.1	21,908.6	31,817.0
Wald (1st stage)	670.1	941.9	4,161.8	6,903.3
Wu-Hausman, p-value	1.72×10^{-7}	0.6623	1.64×10^{-52}	1.88×10^{-73}

(d) idleness rate

Table D.14: Leave-one-out (LOO) Bartik construction robustness test

Notes: This table replicates the main estimates using a slightly different instrument. Recall that in the Bartik-style instrument, local populations are approximated using base-year populations and national-level growth rates. In the LOO specification, when constructing the instrument for a particular PUMA, the national-level growth rate is computed by leaving out that PUMA. Then, the growth rate used in the approximation will be independent of the idiosyncrasies of the PUMA. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	-0.0003	0.0288*	0.0053	0.0134	-0.0100	0.0299*	0.0054	0.0087
	(0.0136)	(0.0149)	(0.0129)	(0.0092)	(0.0143)	(0.0158)	(0.0084)	(0.0072)
2SLS								
Competition (log)	-0.1215**	0.1243**	-0.0901*	0.0281	-0.0981	0.1651**	-0.0524	0.0071
	(0.0600)	(0.0495)	(0.0526)	(0.0284)	(0.0643)	(0.0680)	(0.0695)	(0.0196)
Fit statistics								
Observations	215,378	183,229	133,564	135,788	215,378	183,229	133,564	135,788
F-test (1st stage)	11,245.6	20,227.5	7,741.6	17,428.6	12,018.5	11,380.9	1,975.8	20,463.6
Wald (1st stage)	664.6	1,002.5	438.7	1,049.7	692.8	949.9	149.4	957.3
Wu-Hausman, p-value	0.0224	0.0158	0.0321	0.5224	0.1125	0.0138	0.3366	0.9246

(a) Non-participation rate

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0119	0.0115	0.0027	-0.0106	0.0127	0.0168	0.0002	-0.0055
	(0.0117)	(0.0130)	(0.0134)	(0.0112)	(0.0129)	(0.0140)	(0.0093)	(0.0090)
2SLS								
Competition (log)	-0.1308**	0.0155	0.1452**	-0.0074	-0.1689***	0.0394	0.1740**	-0.0193
	(0.0520)	(0.0354)	(0.0590)	(0.0295)	(0.0552)	(0.0504)	(0.0752)	(0.0218)
Fit statistics								
Observations	178,640	149,872	122,626	128,069	178,640	149,872	122,626	128,069
F-test (1st stage)	9,489.6	17,271.7	7,301.4	16,455.9	10,078.2	9,509.1	1,967.8	19,629.6
Wald (1st stage)	605.1	933.2	408.6	1,012.9	634.0	856.9	143.3	918.9
Wu-Hausman, p-value	0.0014	0.9038	0.0048	0.9063	0.0001	0.6256	0.0092	0.4823

(b) log hours

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0118	0.0244	0.0282	-0.0083	0.0016	0.0208	0.0024	-0.0168
	(0.0306)	(0.0329)	(0.0451)	(0.0335)	(0.0340)	(0.0351)	(0.0294)	(0.0250)
2SLS								
Competition (log)	-0.0515	-0.0837	1.426***	0.1516	-0.1630	0.0163	1.407***	0.0915
	(0.1571)	(0.1109)	(0.1943)	(0.0937)	(0.1688)	(0.1442)	(0.2480)	(0.0716)
Fit statistics								
Observations	178,640	149,872	122,626	128,069	178,640	149,872	122,626	128,069
F-test (1st stage)	9,489.6	17,271.7	7,301.4	16,455.9	10,078.2	9,509.1	1,967.8	19,629.6
Wald (1st stage)	605.1	933.2	408.6	1,012.9	634.0	856.9	143.3	918.9
Wu-Hausman, p-value	0.6009	0.1994	1.56×10^{-20}	0.0303	0.1960	0.9696	1.58×10^{-12}	0.0410

(c) income

competition measure	**baseline**		**with education**	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	-0.0016	-0.0062	-0.0213***	-0.0309***
	(0.0041)	(0.0062)	(0.0039)	(0.0052)
2SLS				
Competition (log)	-0.1188***	-0.0185	-0.1434***	-0.1635***
	(0.0274)	(0.0247)	(0.0122)	(0.0216)
Fit statistics				
Observations	120,231	119,610	120,231	119,610
F-test (1st stage)	3,498.8	7,318.0	15,846.2	9,535.4
Wald (1st stage)	587.8	894.4	2,784.6	1,115.7
Wu-Hausman, p-value	9.83×10^{-7}	0.5871	1.48×10^{-30}	1.15×10^{-14}

(d) idleness rate

Table D.15: "Courtship" robustness test

Notes: In this table we adjust for "courtship" by replacing $\mathcal{P}^s_{x,x'}$ and $P^s_{x,x'}$ with $\mathcal{P}^s_{a+c,a'+c}$ and $P^s_{a+c,a'+c}$, respectively, in the competition meausure (where c stands for courtship). The idea is that if an individual meets a potential marriage partner this year, they expect it to take c years before marriage. That is, "courtship" takes time. In practice, we take c to be 3. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Comp. (long diff.)	-0.0051	0.0299*	0.0040	0.0146	-0.0123	0.0295*	-0.0003	0.0121
	(0.0089)	(0.0153)	(0.0088)	(0.0100)	(0.0092)	(0.0159)	(0.0059)	(0.0078)
2SLS								
Comp. (long diff.)	-0.0708*	0.1041**	-0.0740**	0.0193	-0.0481	0.1548***	-0.1099**	0.0061
	(0.0382)	(0.0415)	(0.0358)	(0.0252)	(0.0434)	(0.0525)	(0.0514)	(0.0175)
Fit statistics								
Observations	215,378	183,229	133,564	135,788	215,378	183,229	133,564	135,788
F-test (1st stage)	11,698.0	32,497.8	7,617.9	26,991.7	10,329.7	19,408.1	2,111.8	30,184.2
Wald (1st stage)	660.7	1,304.0	413.3	1,225.2	685.4	925.9	192.4	1,463.8
Wu-Hausman, p-value	0.0567	0.0229	0.0098	0.8146	0.3517	0.0036	0.0071	0.6786

(a) Non-participation rate

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Comp. (long diff.)	0.0074	0.0158	0.0011	-0.0166	0.0071	0.0257*	-0.0016	-0.0123
	(0.0077)	(0.0137)	(0.0091)	(0.0118)	(0.0083)	(0.0149)	(0.0067)	(0.0094)
2SLS								
Comp. (long diff.)	-0.0958***	0.0126	0.0908**	-0.0108	-0.1191***	0.0055	0.1081**	-0.0199
	(0.0336)	(0.0298)	(0.0396)	(0.0265)	(0.0370)	(0.0404)	(0.0531)	(0.0200)
Fit statistics								
Observations	178,640	149,872	122,626	128,069	178,640	149,872	122,626	128,069
F-test (1st stage)	9,775.3	27,648.1	7,220.1	25,629.0	8,668.2	16,239.3	2,026.0	28,665.9
Wald (1st stage)	589.5	1,227.4	378.5	1,214.7	630.0	829.4	184.7	1,383.7
Wu-Hausman, p-value	0.0004	0.9084	0.0089	0.8043	0.0001	0.5801	0.0176	0.6583

(b) log hours

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Comp. (long diff.)	0.0164	0.0366	0.0389	-0.0029	-0.0048	0.0295	0.0110	-0.0122
	(0.0203)	(0.0338)	(0.0295)	(0.0354)	(0.0221)	(0.0367)	(0.0202)	(0.0272)
2SLS								
Comp. (long diff.)	-0.0208	-0.0797	0.9307***	0.1599**	-0.1050	-0.1375	1.083***	0.0555
	(0.1024)	(0.0901)	(0.1317)	(0.0813)	(0.1176)	(0.1141)	(0.1843)	(0.0639)
Fit statistics								
Observations	178,640	149,872	122,626	128,069	178,640	149,872	122,626	128,069
F-test (1st stage)	9,775.3	27,648.1	7,220.1	25,629.0	8,668.2	16,239.3	2,026.0	28,665.9
Wald (1st stage)	589.5	1,227.4	378.5	1,214.7	630.0	829.4	184.7	1,383.7
Wu-Hausman, p-value	0.6390	0.0960	2.27×10^{-18}	0.0100	0.2559	0.0724	6.58×10^{-15}	0.1433

(c) income

competition measure	baseline		with education	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Comp. (long diff.)	-0.0004	-0.0059*	-0.0051***	-0.0289***
	(0.0011)	(0.0036)	(0.0013)	(0.0040)
2SLS				
Comp. (long diff.)	-0.0187***	-0.0127	-0.0835***	-0.1835***
	(0.0043)	(0.0087)	(0.0062)	(0.0117)
Fit statistics				
Observations	120,231	119,610	120,231	119,610
F-test (1st stage)	8,682.2	20,765.7	6,861.1	14,552.8
Wald (1st stage)	480.3	1,334.7	803.3	233.1
Wu-Hausman, p-value	1.22×10^{-5}	0.4131	4.34×10^{-48}	4.19×10^{-57}

(d) idleness rate

Table D.16: Long difference robustness test

Notes: For this robustness test, instead of using the log competition value as the endogenous regressor (and the log of the approximated competition value as the instrument) in the 2SLS setup, we use the change in competition from a base year (and the change in the approximated competition). Significance codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competitors (log)	-0.0630***	-0.0431	-0.0214	0.0368**	-0.0449**	-0.0281	-0.0164	0.0277**
	(0.0211)	(0.0296)	(0.0185)	(0.0182)	(0.0216)	(0.0313)	(0.0139)	(0.0139)
Eligibles (log)	0.0008	-0.0249*	-0.0077	-0.0155*	0.0072	-0.0247*	-0.0037	-0.0118*
	(0.0110)	(0.0135)	(0.0103)	(0.0085)	(0.0120)	(0.0143)	(0.0067)	(0.0067)
2SLS								
Competitors (log)	0.0026	0.2238*	-0.2336***	-0.1710**	0.0096	0.2671**	-0.1462**	-0.1486**
	(0.1140)	(0.1259)	(0.0871)	(0.0831)	(0.1315)	(0.1161)	(0.0714)	(0.0650)
Eligibles (log)	0.0610	-0.1380**	0.1337**	0.0419	0.0475	-0.1865***	0.0993*	0.0492*
	(0.0608)	(0.0562)	(0.0521)	(0.0361)	(0.0788)	(0.0694)	(0.0532)	(0.0280)
Fit statistics								
Observations	215,378	183,229	133,564	135,788	215,378	183,229	133,564	135,788
F-test (competitors)	21,340.3	18,141.6	16,625.6	15,166.2	36,884.4	30,570.4	22,641.8	16,906.5
F-test (eligibles)	18,696.7	18,834.1	13,208.2	16,280.6	29,115.5	15,217.1	8,299.8	21,128.9
Wald (competitors)	888.4	691.8	631.6	577.8	1,246.3	1,064.8	749.8	605.5
Wald (eligibles)	839.3	771.5	567.0	715.2	1,007.8	812.6	451.7	778.6
Wu-Hausman, p-value	1.99×10^{-5}	0.0239	0.0047	0.0012	0.0001	0.0062	0.0657	0.0021

(a) Non-participation rate

competition measure	baseline		with education	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competitors (log)	0.0012	0.0095	-0.0182**	-0.0571***
	(0.0119)	(0.0143)	(0.0084)	(0.0079)
Eligibles (log)	0.0016	0.0078	0.0258***	0.0508***
	(0.0041)	(0.0058)	(0.0037)	(0.0045)
2SLS				
Competitors (log)	-0.1553***	-0.0554	-0.1160***	-0.2435***
	(0.0532)	(0.1694)	(0.0174)	(0.0156)
Eligibles (log)	0.1226***	0.0337	0.1493***	0.1934***
	(0.0291)	(0.0778)	(0.0097)	(0.0097)
Fit statistics				
Observations	120,231	119,610	120,231	119,610
F-test (competitors)	18,168.0	21,955.4	29,689.6	40,398.5
F-test (eligibles)	6,644.0	14,226.8	15,296.0	26,567.8
Wald (competitors)	1,285.6	923.3	1,153.3	1,478.5
Wald (eligibles)	704.5	901.2	2,502.6	3,906.7
Wu-Hausman, p-value	5.23×10^{-7}	0.8548	2.86×10^{-52}	1.47×10^{-74}

(b) idleness rate

Table D.17: Splitting competition measure into competitors and eligibles

Notes: Significance codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competitors (log)	0.0442^{**}	0.0240	-9.06×10^{-5}	-0.0114	0.0537^{***}	0.0136	-0.0006	0.0015
	(0.0183)	(0.0239)	(0.0192)	(0.0232)	(0.0186)	(0.0251)	(0.0146)	(0.0167)
Eligibles (log)	-0.0087	-0.0134	-0.0024	0.0104	-0.0116	-0.0186	0.0003	0.0090
	(0.0092)	(0.0119)	(0.0108)	(0.0104)	(0.0103)	(0.0128)	(0.0077)	(0.0086)
2SLS								
Competitors (log)	-0.2444^{***}	-0.2820^{***}	-0.1116	-0.1057	-0.2309^{**}	-0.1906^{**}	0.0141	-0.1405^{*}
	(0.0861)	(0.0905)	(0.0915)	(0.0979)	(0.1015)	(0.0852)	(0.0779)	(0.0780)
Eligibles (log)	0.1542^{***}	0.0812^{**}	-0.0366	0.0391	0.1719^{***}	0.0627	-0.0973^{*}	0.0610^{*}
	(0.0483)	(0.0392)	(0.0540)	(0.0415)	(0.0622)	(0.0511)	(0.0564)	(0.0341)
Fit statistics								
Observations	178,640	149,872	122,626	128,069	178,640	149,872	122,626	128,069
F-test (competitors)	18,656.7	15,457.0	15,521.5	14,322.1	31,830.1	25,746.0	21,200.4	15,960.9
F-test (eligibles)	15,786.0	16,113.1	12,204.2	15,394.2	24,254.5	12,885.2	7,727.8	20,060.6
Wald (competitors)	881.7	659.7	610.9	570.2	1,227.6	994.5	720.2	595.0
Wald (eligibles)	782.8	759.9	546.2	705.9	917.3	793.9	436.3	765.7
Wu-Hausman, p-value	0.0011	0.0006	1.11×10^{-6}	0.4233	0.0034	0.0144	4.12×10^{-7}	0.0733

(a) log hours

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competitors (log)	0.1295^{***}	0.1860^{***}	0.0122	-0.1148^{*}	0.1071^{**}	0.0649	0.0476	-0.0604
	(0.0486)	(0.0653)	(0.0679)	(0.0624)	(0.0507)	(0.0700)	(0.0529)	(0.0476)
Eligibles (log)	-0.0111	-0.0350	-0.0398	0.0129	-0.0018	-0.0210	-0.0082	0.0148
	(0.0241)	(0.0290)	(0.0355)	(0.0299)	(0.0276)	(0.0314)	(0.0235)	(0.0239)
2SLS								
Competitors (log)	-0.4512^{*}	-0.9792^{***}	1.013^{***}	-1.395^{***}	-0.4255	-0.9103^{***}	1.004^{***}	-1.130^{***}
	(0.2564)	(0.2612)	(0.2962)	(0.2912)	(0.2943)	(0.2454)	(0.2549)	(0.2178)
Eligibles (log)	0.1859	0.3365^{***}	-1.083^{***}	0.3534^{***}	0.2174	0.3578^{**}	-1.088^{***}	0.3260^{***}
	(0.1492)	(0.1236)	(0.1769)	(0.1324)	(0.1890)	(0.1514)	(0.1816)	(0.1018)
Fit statistics								
Observations	178,640	149,872	122,626	128,069	178,640	149,872	122,626	128,069
F-test (competitors)	18,656.7	15,457.0	15,521.5	14,322.1	31,830.1	25,746.0	21,200.4	15,960.9
F-test (eligibles)	15,786.0	16,113.1	12,204.2	15,394.2	24,254.5	12,885.2	7,727.8	20,060.6
Wald (competitors)	881.7	659.7	610.9	570.2	1,227.6	994.5	720.2	595.0
Wald (eligibles)	782.8	759.9	546.2	705.9	917.3	793.9	436.3	765.7
Wu-Hausman, p-value	0.0022	9.85×10^{-8}	2.75×10^{-21}	5.99×10^{-12}	0.0029	2.28×10^{-7}	1.67×10^{-21}	6.42×10^{-13}

(b) income

Table D.18: Splitting competition measure into competitors and eligibles

Notes: Significance codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	-0.0270	0.0471**	0.0012	0.0168	-0.0278	0.0510**	0.0006	0.0146
	(0.0170)	(0.0207)	(0.0157)	(0.0133)	(0.0190)	(0.0229)	(0.0100)	(0.0105)
2SLS								
Competition (log)	-0.0173	0.1467*	0.0817	0.0817*	-0.0643	0.2082**	0.0276	0.0412
	(0.1120)	(0.0810)	(0.0897)	(0.0468)	(0.1360)	(0.1060)	(0.0666)	(0.0326)
Fit statistics								
Observations	104,725	88,281	61,080	62,751	104,725	88,281	61,080	62,751
F-test (1st stage)	2,770.0	7,991.9	2,000.4	7,701.6	2,364.1	4,910.3	1,499.0	9,403.8
Wald (1st stage)	377.6	572.1	232.5	587.6	333.9	410.9	132.5	550.9
Wu-Hausman, p-value	0.9174	0.1159	0.3030	0.0585	0.7511	0.0707	0.6572	0.2852

(a) Non-participation rate

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0180	0.0630***	0.0319	-0.0132	0.0208	0.0676***	0.0016	-0.0115
	(0.0146)	(0.0181)	(0.0199)	(0.0145)	(0.0171)	(0.0201)	(0.0138)	(0.0118)
2SLS								
Competition (log)	-0.2106**	0.0876	-0.0914	0.0653	-0.3107***	0.0761	-0.1122	0.0423
	(0.0843)	(0.0574)	(0.0974)	(0.0446)	(0.1092)	(0.0786)	(0.0776)	(0.0339)
Fit statistics								
Observations	85,429	71,743	55,633	58,967	85,429	71,743	55,633	58,967
F-test (1st stage)	2,450.3	6,997.9	1,860.1	7,235.4	2,057.6	4,146.9	1,435.6	8,952.3
Wald (1st stage)	362.0	531.8	206.2	573.8	298.8	358.4	122.8	526.4
Wu-Hausman, p-value	0.0036	0.6398	0.1819	0.0548	0.0007	0.9075	0.1081	0.0677

(b) log hours

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0269	0.0955**	0.0580	0.0094	-0.0080	0.0877*	-0.0034	0.0008
	(0.0397)	(0.0462)	(0.0612)	(0.0468)	(0.0458)	(0.0500)	(0.0387)	(0.0373)
2SLS								
Competition (log)	-0.8151***	0.0240	0.3046	0.3695**	-0.9334***	-0.0450	0.1662	0.2790**
	(0.2460)	(0.1570)	(0.3115)	(0.1597)	(0.2978)	(0.2129)	(0.2433)	(0.1190)
Fit statistics								
Observations	85,429	71,743	55,633	58,967	85,429	71,743	55,633	58,967
F-test (1st stage)	2,450.3	6,997.9	1,860.1	7,235.4	2,057.6	4,146.9	1,435.6	8,952.3
Wald (1st stage)	362.0	531.8	206.2	573.8	298.8	358.4	122.8	526.4
Wu-Hausman, p-value	8.2×10^{-5}	0.6028	0.3711	0.0014	0.0005	0.4890	0.4223	0.0006

(c) income

competition measure	baseline		with education	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	-0.0063	-0.0118	-0.0271***	-0.0517***
	(0.0060)	(0.0079)	(0.0053)	(0.0061)
2SLS				
Competition (log)	-0.1062***	0.0007	-0.1421***	-0.2023***
	(0.0349)	(0.0570)	(0.0126)	(0.0148)
Fit statistics				
Observations	60,123	59,832	60,123	59,832
F-test (1st stage)	1,912.4	1,176.8	11,535.6	15,095.9
Wald (1st stage)	286.3	165.8	2,907.0	5,232.7
Wu-Hausman, p-value	0.0021	0.8265	1.69×10^{-22}	1.8×10^{-34}

(d) idleness rate

Table D.19: Restricting analysis to 2006-2013

Notes: Significance codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0162	0.0023	0.0181	0.0230*	0.0118	0.0090	0.0133	0.0173
	(0.0155)	(0.0213)	(0.0145)	(0.0135)	(0.0168)	(0.0212)	(0.0099)	(0.0107)
2SLS								
Competition (log)	-0.1699**	0.3085*	-0.0063	0.1773**	-0.1431	0.4452**	0.0264	0.1138
	(0.0790)	(0.1589)	(0.0670)	(0.0874)	(0.0952)	(0.2079)	(0.0892)	(0.0733)
Fit statistics								
Observations	110,653	94,948	72,484	73,037	110,653	94,948	72,484	73,037
F-test (1st stage)	4,189.5	2,141.1	2,891.3	1,889.7	3,753.4	1,290.7	865.0	1,724.6
Wald (1st stage)	412.5	312.9	180.0	221.9	335.1	191.9	87.31	150.7
Wu-Hausman, p-value	0.0097	0.0142	0.6945	0.0361	0.0652	0.0081	0.8668	0.1211

(a) Non-participation rate

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0116	-0.0002	-0.0080	-0.0045	0.0139	0.0053	0.0052	-0.0078
	(0.0135)	(0.0188)	(0.0158)	(0.0169)	(0.0147)	(0.0194)	(0.0108)	(0.0138)
2SLS								
Competition (log)	0.0280	-0.0542	0.0557	-0.0210	0.0529	-0.0417	0.0983	-0.0467
	(0.0656)	(0.1167)	(0.0843)	(0.0952)	(0.0756)	(0.1543)	(0.1008)	(0.0812)
Fit statistics								
Observations	93,211	78,129	66,993	69,102	93,211	78,129	66,993	69,102
F-test (1st stage)	3,618.7	1,771.5	2,691.3	1,777.2	3,065.0	1,057.6	846.7	1,661.3
Wald (1st stage)	393.9	299.9	171.2	215.7	300.0	171.8	88.65	147.0
Wu-Hausman, p-value	0.7846	0.6209	0.3686	0.8522	0.5880	0.7451	0.2848	0.5976

(b) log hours

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0238	0.0646	-0.0023	0.0221	0.0095	0.0321	0.0318	0.0074
	(0.0350)	(0.0447)	(0.0474)	(0.0419)	(0.0385)	(0.0462)	(0.0323)	(0.0346)
2SLS								
Competition (log)	0.4850***	0.0018	0.7888***	0.1663	0.5072**	-0.1210	0.6098**	0.2772
	(0.1791)	(0.3237)	(0.2470)	(0.2526)	(0.2132)	(0.4124)	(0.3065)	(0.2166)
Fit statistics								
Observations	93,211	78,129	66,993	69,102	93,211	78,129	66,993	69,102
F-test (1st stage)	3,618.7	1,771.5	2,691.3	1,777.2	3,065.0	1,057.6	846.7	1,661.3
Wald (1st stage)	393.9	299.9	171.2	215.7	300.0	171.8	88.65	147.0
Wu-Hausman, p-value	0.0044	0.8178	0.0002	0.5441	0.0104	0.6718	0.0256	0.1735

(c) income

competition measure	baseline		with education	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	0.0027	0.0070	-0.0249***	-0.0464***
	(0.0063)	(0.0090)	(0.0052)	(0.0063)
2SLS				
Competition (log)	-0.1215**	0.1987**	-0.1440***	-0.1855***
	(0.0528)	(0.0869)	(0.0138)	(0.0136)
Fit statistics				
Observations	60,108	59,778	60,108	59,778
F-test (1st stage)	1,011.0	704.4	11,250.2	16,171.5
Wald (1st stage)	171.9	131.4	2,487.2	3,991.0
Wu-Hausman, p-value	0.0076	0.0122	1.6×10^{-23}	1.91×10^{-33}

(d) idleness rate

Table D.20: Restricting analysis to 2014-2021

Notes: Significance codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0093	0.0411^{***}	0.0085	0.0075	0.0062	0.0303^{**}	0.0058	0.0053
	(0.0087)	(0.0111)	(0.0081)	(0.0076)	(0.0094)	(0.0118)	(0.0054)	(0.0060)
2SLS								
Competition (log)	-0.0840^{*}	0.2673^{***}	-0.0004	-0.0162	-0.0999	0.3713^{***}	-0.0013	-0.0241
	(0.0507)	(0.0440)	(0.0418)	(0.0285)	(0.0737)	(0.0807)	(0.0415)	(0.0193)
Fit statistics								
Observations	238,284	225,222	159,106	171,704	238,284	225,222	159,106	171,704
F-test (1st stage)	7,592.2	18,593.9	6,438.0	15,270.4	4,433.9	6,282.1	3,074.7	18,902.1
Wald (1st stage)	846.0	867.8	678.2	828.5	515.5	359.4	347.3	999.0
Wu-Hausman, p-value	0.0332	1.11×10^{-10}	0.7990	0.3041	0.0899	6.97×10^{-8}	0.8379	0.0740

(a) Non-participation rate

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0009	0.0090	-0.0046	-0.0169^{*}	0.0079	0.0124	-0.0039	-0.0165^{**}
	(0.0073)	(0.0092)	(0.0097)	(0.0090)	(0.0081)	(0.0095)	(0.0068)	(0.0074)
2SLS								
Competition (log)	-0.0162	-0.0415	0.1102^{**}	-0.0021	-0.0098	-0.0463	0.1424^{***}	-0.0129
	(0.0387)	(0.0301)	(0.0474)	(0.0298)	(0.0542)	(0.0521)	(0.0465)	(0.0222)
Fit statistics								
Observations	216,367	200,190	148,452	163,447	216,367	200,190	148,452	163,447
F-test (1st stage)	7,073.6	16,921.7	6,030.7	14,475.1	4,129.8	5,627.7	2,916.9	18,092.6
Wald (1st stage)	829.4	835.5	649.9	807.9	514.9	337.5	323.9	987.4
Wu-Hausman, p-value	0.6260	0.0665	0.0045	0.5703	0.7236	0.2398	0.0003	0.8461

(b) log hours

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	-0.0363^{*}	-0.0025	-0.0340	-0.0419^{*}	-0.0335	-0.0092	-0.0312	-0.0514^{***}
	(0.0189)	(0.0242)	(0.0300)	(0.0250)	(0.0212)	(0.0254)	(0.0207)	(0.0194)
2SLS								
Competition (log)	0.0054	-0.2939^{***}	0.5308^{***}	0.0452	0.1040	-0.2974^{**}	0.6695^{***}	-0.0246
	(0.1180)	(0.0890)	(0.1540)	(0.0839)	(0.1670)	(0.1503)	(0.1389)	(0.0653)
Fit statistics								
Observations	216,367	200,190	148,452	163,447	216,367	200,190	148,452	163,447
F-test (1st stage)	7,073.6	16,921.7	6,030.7	14,475.1	4,129.8	5,627.7	2,916.9	18,092.6
Wald (1st stage)	829.4	835.5	649.9	807.9	514.9	337.5	323.9	987.4
Wu-Hausman, p-value	0.6697	4.9×10^{-5}	2.82×10^{-6}	0.2208	0.3237	0.0268	4.4×10^{-9}	0.5964

(c) income

competition measure	baseline		with education	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	0.0014	-0.0009	-0.0204^{***}	-0.0538^{***}
	(0.0041)	(0.0061)	(0.0036)	(0.0045)
2SLS				
Competition (log)	-0.1169^{***}	0.0339	-0.1384^{***}	-0.2163^{***}
	(0.0268)	(0.0236)	(0.0098)	(0.0103)
Fit statistics				
Observations	120,392	120,177	120,392	120,177
F-test (1st stage)	3,817.4	8,576.3	20,908.3	31,646.9
Wald (1st stage)	581.9	916.6	4,169.5	7,369.2
Wu-Hausman, p-value	1.27×10^{-7}	0.0959	3.54×10^{-45}	3.13×10^{-85}

(d) idleness rate

Table D.21: "Unmarried" specification

Notes: The main analysis is repeated except with the "unmarried" population instead of the "single" populations. "Single" individuals have never been married and are not cohabitating with a partner they are not married to. "Unmarried" individuals may have been married before and may be cohabitating with a partner they are not married to. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	-0.0117***	-0.0070	0.0080*	0.0077*	-0.0114**	-0.0057	0.0067**	0.0076**
	(0.0044)	(0.0085)	(0.0047)	(0.0042)	(0.0053)	(0.0098)	(0.0032)	(0.0034)
2SLS								
Competition (log)	-0.0748	0.1676*	0.0050	0.0136	-0.0784	0.2107*	-0.0059	-0.0118
	(0.0613)	(0.0887)	(0.0452)	(0.0353)	(0.0805)	(0.1115)	(0.0628)	(0.0220)
Fit statistics								
Observations	212,064	145,743	132,176	126,516	212,064	145,743	132,176	126,516
F-test (1st stage)	1,182.0	1,703.2	1,336.6	2,157.3	947.0	1,323.8	387.9	3,383.1
Wald (1st stage)	132.8	216.9	140.9	292.5	98.51	155.6	54.88	336.3
Wu-Hausman, p-value	0.2385	0.0116	0.9425	0.8448	0.3506	0.0150	0.8161	0.3159

(a) Non-participation rate

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0083**	0.0184**	0.0009	-0.0057	0.0101**	0.0128	-0.0004	-0.0020
	(0.0040)	(0.0080)	(0.0051)	(0.0057)	(0.0048)	(0.0091)	(0.0037)	(0.0046)
2SLS								
Competition (log)	0.0116	0.0622	0.1241**	0.0137	0.0298	0.0573	0.1041	-0.0231
	(0.0504)	(0.0663)	(0.0532)	(0.0403)	(0.0659)	(0.0887)	(0.0710)	(0.0269)
Fit statistics								
Observations	173,851	110,642	121,349	119,389	173,851	110,642	121,349	119,389
F-test (1st stage)	1,067.0	1,442.2	1,268.3	1,980.4	906.8	1,057.9	370.5	3,202.6
Wald (1st stage)	137.3	193.8	133.4	279.4	105.5	130.9	52.79	323.5
Wu-Hausman, p-value	0.9410	0.4700	0.0080	0.6001	0.7352	0.5799	0.0897	0.3719

(b) log hours

competition measure	**baseline**				**with education**			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0080	0.0666***	0.0204	-0.0056	0.0069	0.0628***	0.0126	-0.0017
	(0.0103)	(0.0194)	(0.0153)	(0.0149)	(0.0126)	(0.0224)	(0.0112)	(0.0123)
2SLS								
Competition (log)	0.1200	-0.0244	0.9619***	0.0885	0.1799	-0.1425	1.199***	0.0195
	(0.1457)	(0.1810)	(0.1851)	(0.1240)	(0.1919)	(0.2297)	(0.2520)	(0.0833)
Fit statistics								
Observations	173,851	110,642	121,349	119,389	173,851	110,642	121,349	119,389
F-test (1st stage)	1,067.0	1,442.2	1,268.3	1,980.4	906.8	1,057.9	370.5	3,202.6
Wald (1st stage)	137.3	193.8	133.4	279.4	105.5	130.9	52.79	323.5
Wu-Hausman, p-value	0.3522	0.5467	9.11×10^{-12}	0.3469	0.2706	0.3046	8.87×10^{-11}	0.7397

(c) income

competition measure	**baseline**		**with education**	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	-0.0012	-0.0043	-0.0165***	-0.0381***
	(0.0031)	(0.0046)	(0.0030)	(0.0035)
2SLS				
Competition (log)	-0.0899***	-0.0760***	-0.1511***	-0.1714***
	(0.0182)	(0.0281)	(0.0105)	(0.0093)
Fit statistics				
Observations	120,208	118,942	120,208	118,942
F-test (1st stage)	4,243.7	2,647.8	12,424.5	23,451.8
Wald (1st stage)	635.7	398.2	2,150.4	4,085.8
Wu-Hausman, p-value	1.67×10^{-8}	0.0101	1.82×10^{-50}	2.95×10^{-73}

(d) idleness rate

Table D.22: "Childless" specification

Notes: The main analysis is repeated except with the "childless" population instead of the "single" populations. "Single" individuals have never been married and are not cohabitating with a partner they are not married to. "Childless" individuals are single but do not have a child in the household (a biological child, a stepchild, or another child). Significance codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	-0.0066	0.0319*	0.0101	0.0047	-0.0116	0.0321*	0.0009	0.0026
	(0.0107)	(0.0167)	(0.0088)	(0.0081)	(0.0121)	(0.0173)	(0.0059)	(0.0066)
2SLS								
Competition (log)	-0.0747	-0.0193	-0.0712*	0.0591***	-0.0750	0.0067	-0.0783*	0.0310*
	(0.0484)	(0.0554)	(0.0373)	(0.0221)	(0.0494)	(0.0598)	(0.0407)	(0.0171)
Fit statistics								
Observations	169,477	125,738	108,491	105,586	169,477	125,738	108,491	105,586
F-test (1st stage)	9,936.8	16,082.6	6,268.7	15,270.0	11,861.9	14,254.2	2,491.4	14,951.1
Wald (1st stage)	994.0	1,084.0	569.4	1,291.6	1,005.2	1,114.0	298.2	1,102.7
Wu-Hausman, p-value	0.0999	0.2334	0.0091	0.0029	0.1440	0.5938	0.0208	0.0537

(a) Non-participation rate

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0088	0.0069	-0.0066	-0.0112	0.0182	-0.0028	-0.0054	-0.0062
	(0.0103)	(0.0159)	(0.0102)	(0.0108)	(0.0121)	(0.0169)	(0.0071)	(0.0087)
2SLS								
Competition (log)	-0.0393	-0.0018	0.1191***	0.0122	-0.0408	-0.0185	0.1048**	0.0026
	(0.0382)	(0.0449)	(0.0427)	(0.0275)	(0.0401)	(0.0495)	(0.0480)	(0.0220)
Fit statistics								
Observations	133,730	96,030	98,939	99,081	133,730	96,030	98,939	99,081
F-test (1st stage)	8,308.7	13,797.3	5,897.7	14,437.8	10,002.4	12,157.2	2,318.7	14,176.8
Wald (1st stage)	898.8	943.8	523.3	1,244.4	925.3	959.4	267.5	1,060.5
Wu-Hausman, p-value	0.1885	0.8180	0.0007	0.3117	0.1232	0.7098	0.0073	0.6369

(b) log hours

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log)	0.0220	0.0252	0.0549*	-0.0263	-0.0008	0.0092	0.0282	-0.0236
	(0.0268)	(0.0386)	(0.0310)	(0.0259)	(0.0307)	(0.0401)	(0.0218)	(0.0207)
2SLS								
Competition (log)	0.0188	0.0117	0.9137***	0.0147	0.0036	0.0055	0.7289***	0.0109
	(0.1168)	(0.1191)	(0.1329)	(0.0731)	(0.1287)	(0.1326)	(0.1442)	(0.0574)
Fit statistics								
Observations	133,730	96,030	98,939	99,081	133,730	96,030	98,939	99,081
F-test (1st stage)	8,308.7	13,797.3	5,897.7	14,437.8	10,002.4	12,157.2	2,318.7	14,176.8
Wald (1st stage)	898.8	943.8	523.3	1,244.4	925.3	959.4	267.5	1,060.5
Wu-Hausman, p-value	0.9734	0.8832	4.02×10^{-15}	0.5092	0.9647	0.9708	7.52×10^{-9}	0.4907

(c) income

competition measure	baseline		with education	
subset	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)
OLS				
Competition (log)	-0.0033	-0.0071	-0.0238***	-0.0378***
	(0.0038)	(0.0049)	(0.0032)	(0.0036)
2SLS				
Competition (log)	-0.1090***	-0.0555***	-0.1372***	-0.1741***
	(0.0298)	(0.0211)	(0.0090)	(0.0104)
Fit statistics				
Observations	112,814	110,785	112,814	110,785
F-test (1st stage)	1,878.8	5,415.8	14,477.2	18,689.0
Wald (1st stage)	360.4	600.6	1,531.2	2,679.6
Wu-Hausman, p-value	8.44×10^{-5}	0.0229	7.86×10^{-43}	3.98×10^{-58}

(d) idleness rate

Table D.23: Whites only

Notes: The main analysis is repeated except with the non-Hispanic white population Significance codes: ***: 0.01, **: 0.05, *: 0.1.

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log) × both in LF	-0.0425	0.0576	-0.0216	0.0082	0.0302	0.0759	-0.0334	0.0016
	(0.0639)	(0.0517)	(0.0542)	(0.0484)	(0.0564)	(0.0582)	(0.0398)	(0.0454)
2SLS								
Competition (log) × both in LF	0.0100	0.1413**	0.0545	0.0130	0.1007	0.1338*	-0.0548	0.0161
	(0.0931)	(0.0647)	(0.0792)	(0.0558)	(0.0673)	(0.0769)	(0.0667)	(0.0596)
Fit statistics								
Observations	215,378	183,229	133,564	135,788	215,378	183,229	133,564	135,788
F-test (1st stage, 1st instr.)	6,088.3	9,752.5	3,967.2	7,766.8	6,430.1	5,566.0	1,402.8	9,122.8
F-test (1st stage, 2nd instr.)	10,122.7	18,830.9	5,623.9	15,120.6	13,651.0	12,997.7	3,019.2	13,947.8
Wald (1st stage, 1st instr.)	388.7	589.9	270.7	588.7	358.7	359.6	138.8	631.7
Wald (1st stage, 2nd instr.)	971.1	1,363.2	417.3	989.1	1,038.0	961.7	279.1	948.3
Wu-Hausman, p-value	0.0472	0.0031	0.0108	0.9721	0.0245	0.0159	0.0989	0.6758

(a) Non-participation rate

competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log) × both in LF	-0.0822	-0.0173	0.1174**	-0.0771	-0.0687	0.0077	0.0939**	-0.0827*
	(0.0511)	(0.0476)	(0.0559)	(0.0475)	(0.0479)	(0.0518)	(0.0438)	(0.0472)
2SLS								
Competition (log) × both in LF	-0.0533	-0.0452	0.1723**	-0.1098*	-0.0769	-0.0300	0.1414*	-0.0800
	(0.0784)	(0.0526)	(0.0862)	(0.0629)	(0.0624)	(0.0604)	(0.0807)	(0.0678)
Fit statistics								
Observations	178,640	149,872	122,626	128,069	178,640	149,872	122,626	128,069
F-test (1st stage, 1st instr.)	5,070.2	8,231.5	3,709.2	7,345.9	5,223.5	4,592.4	1,340.0	8,668.3
F-test (1st stage, 2nd instr.)	8,148.3	15,284.1	5,126.5	14,212.6	10,653.7	10,558.3	2,827.3	13,163.5
Wald (1st stage, 1st instr.)	352.1	538.0	248.9	569.6	317.8	317.4	129.2	609.6
Wald (1st stage, 2nd instr.)	834.0	1,245.9	379.4	953.1	905.7	877.7	259.6	905.7
Wu-Hausman, p-value	0.0043	0.4785	0.0090	0.5899	0.0012	0.4529	0.0136	0.9237

(b) log hours

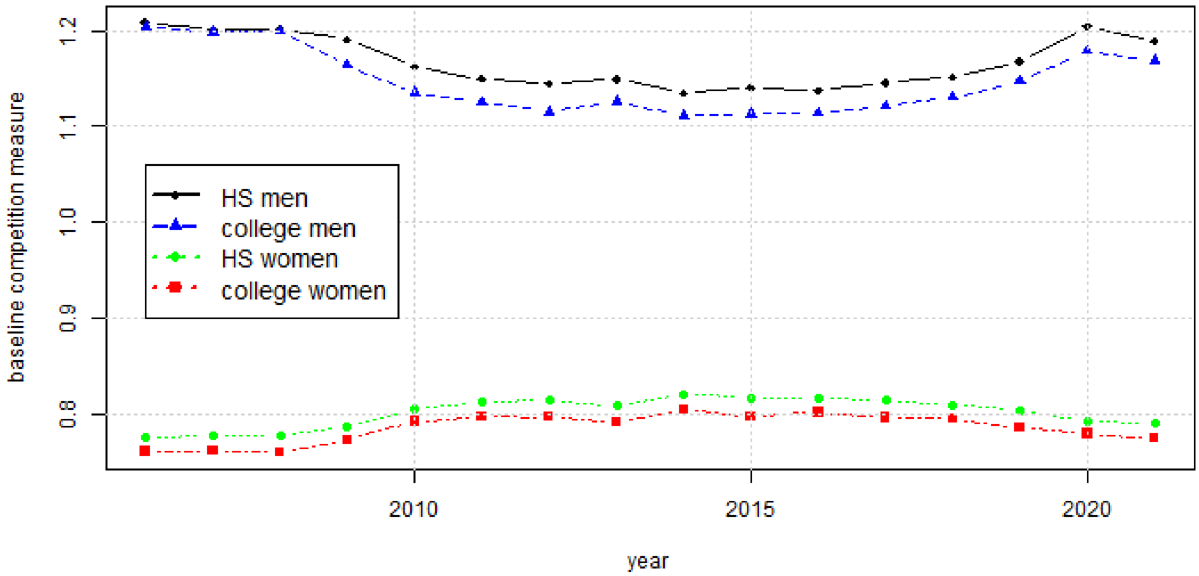
competition measure	baseline				with education			
education	HS		college		HS		college	
sex	*men*	*women*	*men*	*women*	*men*	*women*	*men*	*women*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS								
Competition (log) × both in LF	-0.0081	-0.1118	-0.0553	-0.2674*	0.0882	-0.1525	0.2261	-0.2498*
	(0.1354)	(0.1167)	(0.1864)	(0.1444)	(0.1270)	(0.1348)	(0.1434)	(0.1480)
2SLS								
Competition (log) × both in LF	0.0681	-0.0907	-0.0115	-0.3957**	0.1407	-0.1311	0.0928	-0.3420*
	(0.2041)	(0.1352)	(0.3042)	(0.1746)	(0.1629)	(0.1654)	(0.2779)	(0.1964)
Fit statistics								
Observations	178,640	149,872	122,626	128,069	178,640	149,872	122,626	128,069
F-test (1st stage, 1st instr.)	5,070.2	8,231.5	3,709.2	7,345.9	5,223.5	4,592.4	1,340.0	8,668.3
F-test (1st stage, 2nd instr.)	8,148.3	15,284.1	5,126.5	14,212.6	10,653.7	10,558.3	2,827.3	13,163.5
Wald (1st stage, 1st instr.)	352.1	538.0	248.9	569.6	317.8	317.4	129.2	609.6
Wald (1st stage, 2nd instr.)	834.0	1,245.9	379.4	953.1	905.7	877.7	259.6	905.7
Wu-Hausman, p-value	0.7591	0.1410	5.79×10^{-19}	0.0519	0.7525	0.2907	7.09×10^{-15}	0.2008

(c) income

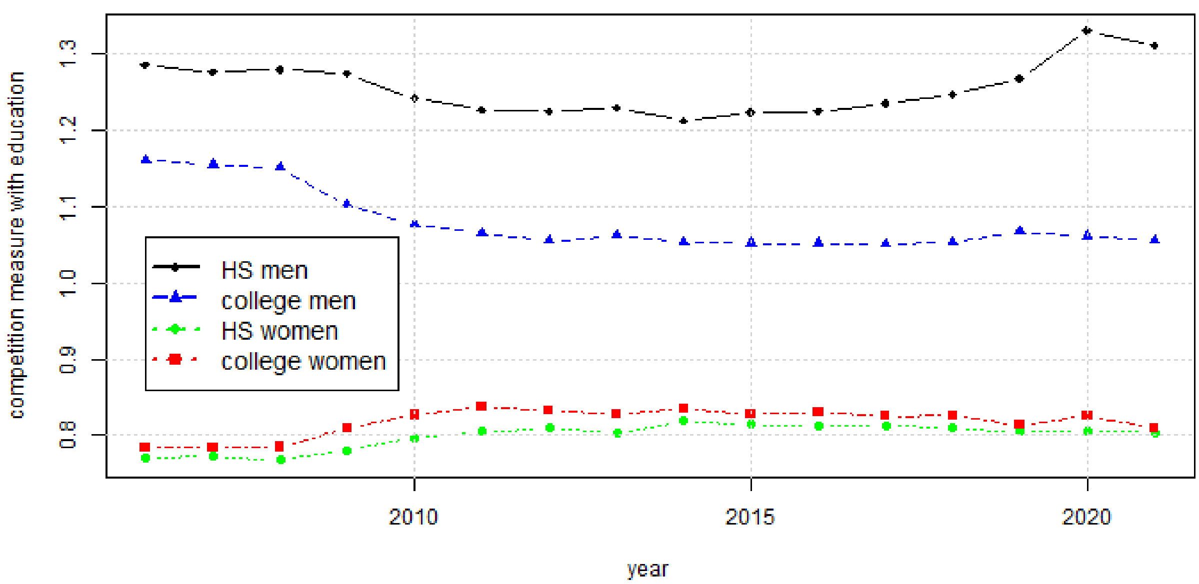
Table D.24: Role of expectations on whether spouse will work

Notes: To gauge this role for men (women), the main analysis is repeated except that competition is interacted with the education-year-location specific fraction of heterosexually married men (women) where both they and their wife (husband) are in the labor force. The interaction term is reported. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

D.7 Competition measure over time



(a) baseline



(b) with education

Figure D.2: Competition measures for 25-39-year-old single adults over time (2006-2021)