Essays in Applied Microeconomics

Benjamin Ferri

A dissertation
submitted to the Faculty of
the department of Economics
in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Boston College
Morrissey College of Arts and Sciences
Graduate School

August 2022

Essays in Applied Microeconomics

Benjamin Ferri

Advisors:

Donald Cox, Ph.D. (Chair)

Christopher Baum, Ph.D.

Hanno Foerster, Ph.D.

Abstract

This dissertation consists of three related chapters. A unifying feature throughout all is a focus on the role of regional earnings distributions, especially at the Commuting Zone level, in driving social and economic behavior. The first chapter examines the role of women's and men's expected earnings, across Commuting Zones, in driving women's and men's location choices (migration). The second chapter, a collaboration with Lia Yin, examines the roles of the upper and lower tails of the earnings distribution in driving crime rates, with a key distinction made between crimes motivated primarily by emotional gain, and those motivated by financial gain. Both chapters one and two use simple structural models, identified by Shift-Share (Bartik) instruments as instrumental variables. The third chapter delves into the history, meaning, and scope of Shift-Share instruments, develops several new variants, and tests them in an application to measuring effects of earnings inequality single parenting rates.

The first chapter, "How Women and Men Choose Where to Live Based on Each Other's Expected Earnings," considers how the distribution of earnings between genders may influence the distribution of the population via internal migration. Might the earnings potential of prospective spouses drive migration choices? Migrants who flock to places with high-earning prospective partners can cause sex ratios to become unbalanced. Shortages of men have been shown to increase rates of single parenting,

and shortages of women to increase crime. Past attempts to answer this question have been limited to brief windows in time, and have lacked causal identification. I build a 7-decade panel of U.S. Commuting Zones from Census and American Community Survey data, computing gender-specific Shift-Share (Bartik) instruments in order to isolate exogenous variation in women's and men's expected earnings. I find that both women and men place at least twice as much priority weight on men's expected earnings as they do on women's, indicating a gender asymmetry in preferences. This asymmetry slightly erodes over time from 1970 to 2019, consistent with a shift in norms. Because women and men prioritize men's earnings over women's by about the same amount, gender differences in earnings play little role in driving sex ratio imbalance. However, women place more weight than men do on the sum of women's and men's earnings, so that the ratio of women to men increases by about 1% per 10% increase in earnings. More balanced sex ratios may follow from policies that reduce overall (gender neutral) inequality, such as between urban and rural areas.

The second chapter, "The Distinct Roles of Poverty and Higher Earnings in Motivating Crime," considers how the two extremes of the earnings distribution bear upon people's propensity to turn to crime. Does inequality lead to more crime? We develop a new model that articulates how Poverty (the lower tail of the earnings distribution) and Earnings (the upper tail) enter into equilibrium crime rates. In our model, individuals in Poverty have less to lose in the context of criminal punishment, so are less averse to committing crimes in general. The presence of high Earnings (therefore things worth stealing) heightens the expected gain to offenders per crime - but specifically in terms of financial gain, not emotional gain. We estimate our model on a comprehensive panel of U.S. Commuting Zones (1980-2016), deploying novel Shift-Share instruments to correct for reverse causality (of crime on the earnings distribution). Corroborating our hypothesis, we find that high Earnings plays a much larger role in driving crimes that yield financial gain to the offender (various forms of theft) than it does for crimes of emotional gain; while Poverty is a driving force equally across both types of crime. In each case, not accounting for reverse causality would underestimate both effects, often by more than double.

The third and final chapter, "Novel Shift-Share Instruments and Their Applications," digs deeper into the topic of Shift-Share (Bartik) instruments, which are vital in both of the earlier chapters. Shift-Share instruments are among the most important tools for causal identification in economics. In this paper, I crystallize main ideas underlying Shift-Share instruments - their core structure, distinctive claim to validity as instruments, history, uses, and wealth of varieties. I argue that the essence of the Shift-Share approach is to decompose the endogenous explanatory variable into an accounting identity with multiple components; preserve that which is most exogenous in the accounting identity, and neutralize that which is most endogenous. Following this framework, I show clearly how several variants in the literature are related. I then develop formulas for several new variants. Particularly, I show how to develop Shift-Share instruments for distribution summaries beyond the mean - the variance, skew, absolute deviation around a central point, and Gini coefficient. As an empirical application that highlights the themes of the paper, I measure the effect of earnings inequality on rates of single parenting in the U.S., comparing results using each of various alternative instruments for the Gini coefficient.

Contents

List of Tables

List of Figures

1 How Women and Men Choose Where to Live Based on Each Other's Expected Earnings

1.1 Introduction

Since Ravenstein (1885), many economists have confirmed that people choose where to live in part on the basis of expected earnings.[1] However, less attention has been paid to gender asymmetry in migration, or to income gained through marriage.[2] Matching with a spouse is similar to matching with an employer in important ways. Both generally depend on physical proximity. And both can be major determinants of income.

Gender differences in location choices can be consequential both for practical and for theoretical reasons. Unbalanced migration can yield unbalanced sex ratios, which are known to have some adverse social effects. Excesses of women can cause marriage rates to decline, and rates of single parenting to increase.[3] Excesses of men can cause crime rates to increase.[4] These effects fall not only on adult decision makers, but also on children, whose welfare adults might not fully internalize. If women and men migrate differently in response to expected earnings, then it may be that policy makers - for example, those interested in altering gender earnings gaps - should consider the secondary effects that would arise due to the shifting of sex ratios.

Migration[5] can also yield rich insights into people's motives. As a significant life choice, the choice of where to live is loaded with information about tastes and

[1] See Greenwood (1997), Kennan and Walker (2011).

[2] I use the words sex and gender as interchangeable and binary, as that is how they are designated in the data.

[3] Angrist (2002) and Charles and Luoh (2010) estimate the effect of sex ratios on marriage, while Harknett (2008) and Dollar (2017) assess the effect on single parenting in particular. Chetty et al. (2014) shows single parenting to be the strongest (amongst plausible candidates) and most robust predictor of socioeconomic immobility.

[4] Edlund et al. (2013), Cameron, Meng and Zhang (2019), Dancygier et al. (2021).

[5] I use the word "migration" as shorthand for, "the choice of where to live, which can include the choice to stay in one's previous location." This is a slightly unusual use of the word, which often restricts to relocation per se.

preferences. Are women more attracted to men's earnings than men are to women's? Becker's formative theory of the household would suppose so.[6] But this supposition has not been tested in the context of location choice, not even for earlier time periods, at least in a way that can include singles.[7] Plus, there have been enormous cultural changes since Becker's time. As I show in Table 1.2, the ratio of female to male college graduates in the US rose from 43/100 in 1970, to 119/100 in 2019.[8] Have 50 years of progress in women's education affected how people choose where to live? Unless women and men already prioritized each other's earnings equally in 1970, it would be surprising if we do not see some change in that direction.

I propose and estimate a model of migration choice, with preferences that may vary by gender. I hypothesize that women and men have Symmetric Preferences. This is the idea that women place similar weight (on average) on men's expected earnings, in deciding where to live, as men do on women's expected earnings.[9] Under Symmetric Preferences, gender differences in migration choices might arise from differences in opportunities between women and men, but not from different preferences. A competing hypothesis of Asymmetric Preferences holds that women and men react differently to underlying opportunities. In particular, it is supposed that women are more reactive to men's expected earnings than men are to women's.

Edlund (2005) develops a compelling theory of migration choice in which women and men have asymmetric preferences. As an equilibrium result arising from underlying gender differences in attitudes towards marriage, Edlund predicts that women would migrate disproportionately to areas in which men's earnings are highest. She also confirms this result empirically. Edlund (2005) is novel in exploring the implications of asymmetric preferences between women and men for migration and sex

[6] See Becker (1973).

[7] Compton and Pollak (2007), Blackburn (2010) and others have studied the joint location decisions of couples with respect to each others' earnings, but this approach cannot include singles.

[8] See also Goldin, Katz and Kuziemko (2006) and Zhang (2021).

[9] More precisely, the amount of weight placed on opposite *relative to* own gender earnings should be equal across genders.

ratios. Relative to this as a starting point, I make three major innovations. First, I write a more detailed model that estimates distinct migration elasticities for women and men, amounting to an empirical test of whether women's and men's preferences are actually Asymmetric.[10] Second, Edlund does not account for reverse causality (of population on earnings).[11] I use shift-share (Bartik) instruments to isolate exogenous variation in women's and men's expected earnings. Third, I use a panel spanning 1950-2019, rather than a cross section. This enables regional fixed effects and year by region effects to absorb unobserved information, which is also important for causal identification. The panel also enables investigation of change over time, which is essential for assessing the impact of cultural change on the preferences in question.

I build a panel of all 722 Commuting Zones of the US mainland states (48+DC), over 7 time periods from 1950 to 2019. The raw data are sourced from IPUMS USA 1% microsamples, and Census and American Community Survey summaries at the county level.[12] Harnessing the detailed information in the microsamples, in combination with the full coverage of the county level summaries, I construct parallel measures for a large variety of demographic types. Measures include, principally, population counts and expected earnings. Demographic types include genders, by age groups, by education levels, by employment status, by marital status, and so on.

The core of the model is a standard discrete choice problem. Each individual person in each time period chooses the Commuting Zone (CZ) with the most desirable bundle of characteristics. These characteristics include women's expected earnings and men's expected earnings. The population count is the sum of those who have

[10]Edlund's empirical specification cannot directly test this; it is rather baked in to the model as an assumption.

[11]The estimation interprets all differences across cities as movements along supply (of population or labor), that is, people moving to where the money is. This is misleading if some differences represent rather movements along demand. For example, men's earnings may be higher in a city with fewer men present, because men's skills are relatively scarce there. Moreover, the calculation of women's (men's) average earnings divides by the number of women (men), sometimes including those who are not there to work.

[12]Ruggles et al. (2020)

moved into the CZ and those who have stayed in the CZ, so is exactly equal to the number who have chosen it as a location. This captures the average level of utility the CZ offers to people of any given demographic type. To account for inertial frictions - such as moving costs, familiarity and social ties - I estimate and subtract the effects of preferences to live in the state in which one was born. I then also impose one lag serial autocorrelation on the remaining portion of the implied average utility offered by each CZ.

My hypothesis of Symmetric Preferences supposes that women place a similar value as men do on earning directly, as opposed to earning through a spouse. The alternative supposes that women are more inclined to earn through a spouse (as opposed to earning directly) than men are.[13] This notion of Asymmetric Preferences is consistent with experimental results,[14] and theories based on evolutionary biology.[15] It would follow that both men and women will respond primarily to men's expected earnings, regardless of the state of women's opportunities.

Under Asymmetric Preferences, if high skilled men live primarily in major cities, women may be incentivized to move disproportionately to major cities, even though there are already more women than men there. This is because the expected benefit of a less certain match with a higher earning man can exceed that of a more certain match with a lower earning man. Gautier, Svarer and Teulings (2010) present evidence that cities serve as more efficient marriage matching markets than other areas do. Chiappori, Salanié and Weiss (2017) find that for Americans born from 1943-1972, men over time became become more inclined to marry women of similar education as themselves, and vice versa. In other words, more towards the past, women's

[13]Under Asymmetric Preferences, men may prefer earning through a spouse to earning directly, but just to a lesser extent than women do. Or, women may prefer earning directly to earning through a spouse, but to a lesser extent than men. It is simply a negation of the idea that women and men evaluate opportunities equally.

[14]Fisman et al. (2006) and Ong and Wang (2015) use speed-dating and online dating experiments, respectively. Attanasio and Kaufmann (2017) use self-reported expectations in manner similar to a vignette experiment. See Croson and Gneezy (2009) for a general review of experimental results of gender differences in preferences.

[15]See Edlund (2005) and Saint-Paul (2015).

education had less effect on their marital match quality. Towards the present, the marriage market became increasingly competitive amongst women on the basis of education. Migration choice may be an additional strategy by which people compete for favorable marriage matches.

This paper's primary contribution is to the literature on the interrelationship between marriage search and migration.[16] In particular, I tackle the question of whether there is gender asymmetry in migration priorities with respect to expected earnings, in a way that includes singles. Much literature[17] has found that married women tend to be Tied Movers, meaning that their husbands' careers play a larger role in their migration choices than their own careers do. Though consistent with preference asymmetry between women and men, this approach is limited to married couples, and therefore cannot explain sex ratio imbalance. Only Edlund (2005) and Kröhnert and Vollmer (2012) have taken an approach similar to mine, but neither did so in the context of an explicit model with preferences, nor with causal identification, nor with change over time.

More broadly, this paper contributes to the literature on the determinants of gender differences in society, from an unusual angle. Many studies[18] focus on the role of external factors, such as discrimination, and skill differences between women and men, in driving overall gender differences in earnings opportunities. I rather investigate internal factors (preferences), by measuring how women and men choose over the set of opportunities available to each of them, in each time period. Similarly, Bertrand, Pan and Kamenica (2013) examine people's choices of whether to marry or divorce, as revealed preference with respect to relative earnings within the household. They find that unions in which the woman earns more than the man are less likely

[16] See Jang, Casterline and Snyder (2014), Compton and Pollak (2014), Weiss, Yi and Zhang (2013).

[17] Blackburn (2010), Løken, Lommerud and Lundberg (2013), Compton and Pollak (2007), Brandén and Haandrikman (2019), Abraham, Bähr and Trappmann (2019), Gubhaju and De Jong (2009).

[18] See Goldin (2014) for a review.

to form and survive, indicating that men tend to be averse to marrying women who earn more than - or women tend to be averse to marrying men who earn less than - themselves. Attanasio and Kaufmann (2017) use the choice of whether to enroll in college as revealed preference, finding that women place relatively more weight on their expected marriage outcomes, while men place relatively more on their expected work outcomes. This paper uses the choice of where to live as revealed preference, which has the unique advantage of enabling investigation of change over a very long span of time.

1.2 Model

My hypothesis of Symmetric Preferences is a prediction about women's and men's preferences with respect to women's and men's expected earnings. I use migration choices to reveal these preferences. To that end, the model serves two main purposes. First, it clarifies how and why migration choices map to preferences. In particular, each (women's, men's) preference parameter with respect to (women's, men's) expected earnings will (in a special case) be exactly equal to a corresponding elasticity of net migration. Second, the model articulates potential confounds that may arise - both in measuring elasticities of net migration and in interpreting them to reveal preferences - and responds to each confound.

1.2.1 Migration Choice

The starting point of the model is a discrete choice problem over the set of location choice options.[19] I take this set of options to be the 722 Commuting Zones (CZs) of the US mainland states (48+DC).[20] Each individual person i, in each time

[19]My migration choice equation is similar to the labor supply equation of Diamond (2016). Migration choice (location choice) is equivalent to labor supply in the sense that each person is a unit of labor quantity. People seek to go where the money is. Therefore the quantity of labor (population count) is an increasing function of the price of labor (expected earnings).

[20]Of course, this limits the analysis to the choice of CZ conditional on the choice to live in the US, and to the set of people who indeed (have chosen to) live in the US in any given time period. The exclusion of Alaska and Hawaii is not essential, though more in line with convention.

period t, chooses to live in the CZ z that presents the highest utility, $u_{i,t,z}$. $u_{i,t,z}$ has three components. The first component, $u_{j,t}(x_{t,z})$, is a common average utility function shared by all members of i's demographic group, j, in evaluating the objective characteristics ($x_{t,z}$) of location options z, such as the expected earnings in each. The second component, $\beta^{b}_{j,t} b_{i,z}$, is the utility value of living in one's state of birth: $b_{i,z}$ is dummy (binary) variable, indicating whether individual i was born in the same state as location option z. The third component, $\varepsilon_{i,t,z}$, is an idiosyncratic remainder term, representing everything else that would make location option z more or less (relative to i's demographic group, j) attractive to i as an individual. That is, the discrete choice problem is,

$$\max_{z} \left\{ u_{i,t,z} = u_{j,t}(x_{t,z}) + \beta^{b}_{j,t} b_{i,z} + \varepsilon_{i,t,z} \right\} \tag{1}$$

where i the individual person, t is the time period, z is any given location option (Commuting Zone in which to live), j is i's demographic type (such as non-institutionalized heterosexual woman in the age range of 18-64),[21] u is expected utility (the component of utility that depends on location), x is the vector of relevant location characteristics (such as expected earnings and weather, that people would evaluate in a common average way in deciding where to live), $b_{i,z}$ indicates whether location option z is in person i's state of birth, and ε is an idiosyncratic (unique to each individual i,z pair) remainder term.

The model interprets *all* observed addresses as migration *choices*, even if the individual did not move. Everyone has the *opportunity* to move, and implicitly considers moving to each location option z, in each time period. Those who choose to stay in the same place z as in a previous time period have made the determination that this place z continues to offer them higher expected utility than all others. Inertial frictions, such as moving costs, familiarity and social connections, (these all are part

[21] j in principle can restrict to any subset of the population.

of the utility, $u_{i,t,z}$) will of course tend to favor places z that the person i has already lived in before. Because inertial frictions are idiosyncratic (that is, specific to each i,z pair), they can play no role in $u_{j,t}(x_{t,z})$, by definition. Rather, the frictions are shared between $\beta^b_{j,t} b_{i,z}$ and $\varepsilon_{i,t,z}$. People will tend to prefer ($\beta^b_{j,t} > 0$) to live in the state in which they were born ($b_{i,z} = 1$) *because of* familiarity, social ties, and typically lower moving costs. Beyond the extent to which the frictions are captured in $b_{i,z}$, they are captured instead in $\varepsilon_{i,t,z}$.[22] I represent this remaining portion by assuming that $\varepsilon_{i,t,z}$ is (one lag) autocorrelated over time, with an autocorrelation coefficient specific to each j group.

Following McFadden (1973), by assumption that $\varepsilon_{i,t,z}$ is EV1 distributed with respect to the choice options z, equation (1) translates into logit choice probabilities. That is,

$$N_{j,t,z} = \sum_i \frac{exp(u_{j,t}(x_{t,z}) + \beta^b_{j,t} b_{i,z})}{\sum_{\mathcal{z}} exp(u_{j,t}(x_{t,\mathcal{z}}) + \beta^b_{j,t} b_{i,\mathcal{z}})} \tag{2}$$

where z, $\mathcal{z}$ are choice options (Commuting Zones), $N_{j,t,z}$ is the population count of people of type j who (have chosen to) live in each location z in time t, and $N_{j,t}$ is the national population count. Maximum Likelihood estimation of (2) recovers estimates $\hat{\beta}^b_{j,t}$ of $\beta^b_{j,t}$ for each j,t, and also scalar values $\hat{u}_{j,t,z}$ representing the output of the function $u_{j,t}(x_{t,z})$ in each j,t,z cell. $\hat{u}_{j,t,z}$ contains the autocorrelation discussed in the previous paragraph. Plugging the estimates back in to (2) and rearranging,

$$\hat{u}_{j,t,z} = n_{j,t,z} - log(exp(\hat{\beta}^b_{j,t}) b_{j,t,z} + 1 - b_{j,t,z}) + \bar{u}_{j,t} \tag{3}$$

$$n_{j,t,z} = log(N_{j,t,z}/N_{j,t}) \ , \ b_{j,t,z} = \sum_i b_{i,z}/N_{j,t} \ , \ \bar{u}_{j,t} = log(\sum_{\mathcal{z}} exp(\hat{u}_{j,t,\mathcal{z}} + \beta^b_{j,t} b_{i,\mathcal{z}}))$$

$$\hat{u}_{j,t,z} = u_{j,t}(x_{t,z}) + \tilde{\varepsilon}_{j,t,z} \tag{4}$$

[22]Similarly, $\varepsilon_{i,t,z}$ also captures personal (specific to i) factors that result in specific job connections, beyond the average effects delivered in $u_{j,t}(x_{t,z})$. For example, in so far as a woman i is more likely to obtain a desirable job in a CZ z in which the expected earnings of women as a group j are higher, this is represented in $u_{j,t}(x_{t,z})$. Individual factors that would not be captured in the group j average, such the individual person's career networking, are instead in $\varepsilon_{i,t,z}$.

where $\tilde{\varepsilon}_{j,t,z}$ represents the autocorrelation over time of $\hat{u}_{j,t,z}$. Equation (3) shows that the average utility ($\hat{u}_{j,t,z}$) that the model recovers for each j,t,z cell is increasing one for one in the log population count ($n_{j,t,z}$) of people who (have chosen to) live in each.[23] This is similar to the idiomatic idea that people "vote with their feet." Equation (4) represents that which remains to be estimated. Observed are the output $\hat{u}_{j,t,z}$ of the function per each j,t,z cell, and potential arguments $x_{t,z}$, such as expected earnings. The functional form, and exact set of relevant arguments, need to be assumed.

1.2.2 Expected Earnings

As arguments $x_{t,z}$ of average utility, I am interested primarily in expected earnings. Higher expected earnings (per location) may be desirable for multiple reasons - to fund consumption, for savings and security, or proximity to opportunities. Even the taxed portion of earnings may be desirable (per location), because local or state taxes are generally necessary to fund local amenities. Rather than modeling these factors separately, I assume that people on average (in deciding where to live take) take expected earnings itself as the main choice factor of interest, representing in sum any underlying reasons for expected earnings to desirable. This may be considered a "behavioral" assumption: people do not have immense powers of prediction pertaining to various highly specific ex post outcomes, but rather operate more on a level of heuristics.[24] Expected earnings (per location), being salient, well-known and visible, are particularly realistic as a choice factor in this sense.

Especially, I am interested in distinguishing between the expected earnings of women, and the expected earnings of men, as choice factors. For women, the expected earnings of women are those that they may expect to access directly (that is, via their

[23] Because $\bar{u}_{j,t}$ is a sum over *all* choice options (z), it is (very nearly) constant over z and i, making it an (unimportant) intercept term. The term $log(exp(\hat{\beta}^{b}_{j,t})b_{j,t,z} + 1 - b_{j,t,z})$ does of course yield some differences between $\hat{u}_{j,t,z}$ and $n_{j,t,z}$.

[24] Similarly, I operate under a realistic set of assumptions pertaining to dynamism, with the result that the lifetime utility arising from location choice z in time period t is fully characterized the flow utility arising from location choice z in time period t. See Appendix section 1.9.5 for details.

own work), and the expected earnings of men are those that they may expect to access through a (current or future) spouse.[25] However, expected access to the opposite gender's earnings may be influenced by the sex ratio. A favorable sex ratio may confer on one better chances of securing a marriage match, ability to select a match of above average quality, or bargaining power. I assume that the combined effects of these factors are captured in a single parameter, σ. That is, women's expected access to men's earnings (the component of expected access that is dependent on the sex ratio) is $(\mathrm{N}_{m,\mathrm{z}}/\mathrm{N}_{w,\mathrm{z}})^{\sigma}$, where w and m denote women and men respectively (so $\mathrm{N}_{m,\mathrm{z}}/\mathrm{N}_{w,\mathrm{z}}$ is the population ratio of men to women), and $\sigma \geq 0$. Because the sex ratio is highly endogenous in this model, σ has to be calibrated rather than estimated. $\sigma = 0$ would mean that expected access is constant with respect the sex ratio, or, that people on average are inattentive to the sex ratio in deciding where to live. I consider this to be the most realistic supposition - that the average person does not think about the sex ratio in deciding where to live. I therefore estimate assuming $\sigma = 0$, and also, as a robustness check, assuming a positive value of σ.[26]

In conjunction with women's and men's expected earnings as arguments of average utility, two other arguments are essential in theory. First, expected leisure time is inevitably in tradeoff with (one's own) expected earnings. In so far as higher earnings coincide with more time and effort working, they naturally would coincide also with less time for everything besides paid work. I refer to the sum of time spent on all activities besides paid work as leisure (though many such activities needn't be leisurely in a colloquial sense), L. Second, any number of non-pecuniary factors (weather, retail opportunities, crime, culture, schools, parks) may make some locations more desirable to live in than others to people on average. Inevitably, at least some such miscellaneous factors, collectively called amenities, A, must be unobserved to the researcher.

[25] I adjust all data used in the model estimation to capture heterosexuals only.

[26] For this positive value, I use the observed elasticity of marriage rates with respect to sex ratios. See Appendix section 1.9.4 for details.

1.2.3 Utility

I assume that the average utility function $\mathrm{u_{j,t}(x_{t,z})}$ is Cobb-Douglas with respect to the four arguments discussed the previous three paragraphs (expected leisure time, L, expected earnings of women and men, Y_w and Y_m, and amenities, A), with parameters that are specific to each demographic group j,[27] and may change over time t. This is, for women and men (w and m) respectively,

$$\mathrm{u}_{\mathrm{j}w,\mathrm{t}}(\mathrm{x_{t,z}}) = \bar{\mathrm{U}}_{\mathrm{j}w,\mathrm{t}}(\mathrm{L}_{\mathrm{j}w,\mathrm{t,z}})^{\beta^l_{\mathrm{j}w,\mathrm{t}}}(\mathring{\mathrm{Y}}_{\mathrm{j}w,\mathrm{t,z}})^{\beta^y_{\mathrm{j}w,\mathrm{t}}}(\mathring{\mathrm{Y}}_{\mathrm{j}m,\mathrm{t,z}}(\mathrm{N}_{\mathrm{j}m,\mathrm{t,z}}/\mathrm{N}_{\mathrm{j}w,\mathrm{t,z}})^{\sigma})^{\beta^y_{\mathrm{j}w,\mathrm{t}}-\beta^{y'}_{\mathrm{j}w,\mathrm{t}}}(\mathrm{A_z})^{\beta^a_{\mathrm{j}w,\mathrm{t}}}$$
$$\mathrm{u}_{\mathrm{j}m,\mathrm{t}}(\mathrm{x_{t,z}}) = \bar{\mathrm{U}}_{\mathrm{j}m,\mathrm{t}}(\mathrm{L}_{\mathrm{j}m,\mathrm{t,z}})^{\beta^l_{\mathrm{j}m,\mathrm{t}}}(\mathring{\mathrm{Y}}_{\mathrm{j}m,\mathrm{t,z}})^{\beta^y_{\mathrm{j}m,\mathrm{t}}}(\mathring{\mathrm{Y}}_{\mathrm{j}w,\mathrm{t,z}}(\mathrm{N}_{\mathrm{j}w,\mathrm{t,z}}/\mathrm{N}_{m\mathrm{j},\mathrm{t,z}})^{\sigma})^{\beta^y_{\mathrm{j}m,\mathrm{t}}-\beta^{y'}_{\mathrm{j}m,\mathrm{t}}}(\mathrm{A_z})^{\beta^a_{\mathrm{j}m,\mathrm{t}}}$$

For readability, for the remainder of this section, I suppress j,t subscripts. As such, the above equation set becomes rather,

$$\begin{aligned}
\mathrm{u}_{w,\mathrm{z}} &= \bar{\mathrm{U}}_w(\mathrm{L}_{w,\mathrm{z}})^{\beta^l_w}(\mathring{\mathrm{Y}}_{w,\mathrm{z}})^{\beta^y_w}(\mathring{\mathrm{Y}}_{m,\mathrm{z}}(\mathrm{N}_{m,\mathrm{z}}/\mathrm{N}_{w,\mathrm{z}})^{\sigma})^{\beta^y_w-\beta^{y'}_w}(\mathrm{A_z})^{\beta^a_w} \\
\mathrm{u}_{m,\mathrm{z}} &= \bar{\mathrm{U}}_m(\mathrm{L}_{m,\mathrm{z}})^{\beta^l_m}(\mathring{\mathrm{Y}}_{m,\mathrm{z}})^{\beta^y_m}(\mathring{\mathrm{Y}}_{w,\mathrm{z}}(\mathrm{N}_{w,\mathrm{z}}/\mathrm{N}_{m,\mathrm{z}})^{\sigma})^{\beta^y_m-\beta^{y'}_m}(\mathrm{A_z})^{\beta^a_m}
\end{aligned} \tag{5}$$

Expected earnings $\mathring{\mathrm{Y}}_{w,\mathrm{z}}$ and $\mathring{\mathrm{Y}}_{m,\mathrm{z}}$ are adjusted by federal tax and local cost of living.[28] The ring symbol over the letter denotes adjustment by cost of living. Because $\mathring{\mathrm{Y}}_{w,\mathrm{z}}$ are the earnings that women (of subgroup j) would on average expect to acquire directly (via their own work) conditional on the choice to live in location option z (in time period t), the utility value (β^y_w) of $\mathring{\mathrm{Y}}_{w,\mathrm{z}}$ to women is simply the utility value of earnings (y) per se. The utility value of men's expected earnings to women is also the value of earnings (β^y_w), except that women have less certainty and control over men's (women's expected spouses') earnings than they do over their own earnings. The parameter $\beta^{y'}_w$ represents this wedge in certainty and control. That women value men's earnings at all, of course assumes that earnings are shared between women and

[27] j can denote any set of demographic restrictions, including gender. When written with a gender indicator also, such as jw, the j stands for any other restrictions, besides gender (such as age range).

[28] See Appendix section 1.9.6 and 1.9.4 for details. $\bar{\mathrm{Y}}$ is national average pre-tax earnings, $\mathrm{Y_z}$ is CZ average post-tax earnings, and $\mathring{\mathrm{Y}}_\mathrm{z}$ is $\mathrm{Y_z}$ adjusted by cost of living in z.

men: a hypothetical case of $\beta^{y'} = 0$ may coincide with perfect sharing.

The functional form of (5) enables the simplest plausible interpretation of the distinction between preferences and opportunities. That the utility is multiplicatively separable in women's and men expected earnings yields the feature that the *overall (national) level* of women's expected earnings (opportunities) relative to men's has no bearing on the arguments ($\mathring{\mathrm{Y}}_{w,\mathrm{z}}$ and $\mathring{\mathrm{Y}}_{m,\mathrm{z}}$), which are meant rather to represent *relative* opportunities across location options z. Imagine for example that all women's earnings everywhere were tripled, and men's held constant: this would strike (5) as,

$$\begin{aligned}
\mathrm{u}_{w,\mathrm{z}} &= \bar{\mathrm{U}}_w(\mathrm{L}_{w,\mathrm{z}})^{\beta^l_w}(3\mathring{\mathrm{Y}}_{w,\mathrm{z}})^{\beta^y_w}(\mathring{\mathrm{Y}}_{m,\mathrm{z}}(\mathrm{N}_{m,\mathrm{z}}/\mathrm{N}_{w,\mathrm{z}})^{\sigma})^{\beta^y_w-\beta^{y'}_w}(\mathrm{A}_\mathrm{z})^{\beta^a_w} \\
&= 3^{\beta^y_w}\bar{\mathrm{U}}_w(\mathrm{L}_{w,\mathrm{z}})^{\beta^l_w}(\mathring{\mathrm{Y}}_{w,\mathrm{z}})^{\beta^y_w}(\mathring{\mathrm{Y}}_{m,\mathrm{z}}(\mathrm{N}_{m,\mathrm{z}}/\mathrm{N}_{w,\mathrm{z}})^{\sigma})^{\beta^y_w-\beta^{y'}_w}(\mathrm{A}_\mathrm{z})^{\beta^a_w} \qquad (6)\\
&= \bar{\mathrm{U}}_w(\mathrm{L}_{w,\mathrm{z}})^{\beta^l_w}(\mathring{\mathrm{Y}}_{w,\mathrm{z}})^{\beta^y_w}(\mathring{\mathrm{Y}}_{w,\mathrm{z}}(\mathrm{N}_{m,\mathrm{z}}/\mathrm{N}_{w,\mathrm{z}})^{\sigma})^{\beta^y_w-\beta^{y'}_w}(\mathrm{A}_\mathrm{z})^{\beta^a_w}
\end{aligned}$$

$\bar{\mathrm{U}}_w$ is an arbitrary scalar, absorbing the product of all variables that do not vary over location (z). Because utility is immune to monotonic transformation, it follows that the value of $\bar{\mathrm{U}}_w$ has no importance at all, and $\bar{\mathrm{U}}_w$ can absorb any other scalar (such as $3^{\beta^y_w}$). Therefore - if this form of utility is true, then - changes in the overall level of women's opportunities relative to men's (such as between 1980 and 2010) *cannot* change migration behavior via the arguments $\mathring{\mathrm{Y}}$, lacking any meaningful impact on *relative* opportunities across locations. Rather, migration behavior would change *only* if the overall national change in opportunities were to coincide with a change in preferences (β) as well.[29]

To simplify further in focusing on women's and men's expected earnings ($\mathring{\mathrm{Y}}_{w,\mathrm{z}}$ and $\mathring{\mathrm{Y}}_{m,\mathrm{z}}$), I recast the role of expected leisure time in terms of its elasticity with respect to expected earnings. Suppose that expected earnings are the product of expected hours of work, H_z, and expected wages, $\mathring{\mathrm{W}}_\mathrm{z}$, and that expected leisure time

[29]This of course represents merely one of many plausible interpretations of the distinction between preferences and opportunities, chosen for maximal simplicity.

is a negative of hours of work. That is,

$$\begin{aligned} \mathring{\mathrm{Y}}_{w,\mathrm{z}} = \mathrm{H}_{w,\mathrm{z}}\mathring{\mathrm{W}}_{w,\mathrm{z}} \ , \ \mathring{\mathrm{Y}}_{m,\mathrm{z}} = \mathrm{H}_{m,\mathrm{z}}\mathring{\mathrm{W}}_{m,\mathrm{z}} \\ \mathrm{L}_{w,\mathrm{z}} = 2\bar{\mathrm{H}}_{w} - \mathrm{H}_{w,\mathrm{z}} \ , \ \mathrm{L}_{w,\mathrm{z}} = 2\bar{\mathrm{H}}_{m} - \mathrm{H}_{m,\mathrm{z}} \end{aligned} \tag{7}$$

where $\bar{\mathrm{H}}$ denotes national average hours of work, typically about 1,500 hours/year. Functionally, I assume that a total of $2\bar{\mathrm{H}}$ time is allotted between expected working time and expected leisure, resulting in the above expression. The elasticity of expected leisure with respect to expected earnings is,

$$\frac{\mathrm{Y}_{\mathrm{z}}}{\mathrm{L}_{\mathrm{z}}}\frac{d\ \mathrm{L}_{\mathrm{z}}}{d\ \mathrm{Y}_{\mathrm{z}}} = \frac{d\ log(\mathrm{L}_{\mathrm{z}})}{d\ \mathrm{Y}_{\mathrm{z}}}\mathrm{Y}_{\mathrm{z}} = \frac{d\ log(2\bar{\mathrm{H}} - \mathrm{Y}_{\mathrm{z}}/\mathrm{W}_{\mathrm{z}})}{d\ \mathrm{Y}_{\mathrm{z}}}\mathrm{Y}_{\mathrm{z}} = \frac{-1}{2\bar{\mathrm{H}}\mathrm{W}_{\mathrm{z}} - \mathrm{Y}_{\mathrm{z}}}\mathrm{Y}_{\mathrm{z}} \tag{8}$$

At average values, $\bar{\mathrm{H}}\mathrm{W}_{\mathrm{z}} = \bar{\mathrm{Y}}$, so the expression above is equal to exactly -1. That is to say, on average, a % increase in expected earnings coincides with an equal % decrease in expected leisure. As such, I allow the expected earnings term in the utility to subsume the role of expected leisure. That is,

$$\begin{aligned} \mathrm{u}_{w,\mathrm{z}} = \bar{\mathrm{U}}_{w}(\mathring{\mathrm{Y}}_{w,\mathrm{z}})^{\left[\beta^{y}_{w} - \tilde{\beta}^{l}_{w}\right]}(\mathring{\mathrm{Y}}_{m,\mathrm{z}}(\mathrm{N}_{m,\mathrm{z}}/\mathrm{N}_{w,\mathrm{z}})^{\sigma})^{\beta^{y}_{w} - \beta^{y\prime}_{w}}(\mathrm{A}_{\mathrm{z}})^{\beta^{a}_{w}} \\ \mathrm{u}_{m,\mathrm{z}} = \bar{\mathrm{U}}_{m}(\mathring{\mathrm{Y}}_{m,\mathrm{z}})^{\left[\beta^{y}_{m} - \tilde{\beta}^{l}_{m}\right]}(\mathring{\mathrm{Y}}_{w,\mathrm{z}}(\mathrm{N}_{w,\mathrm{z}}/\mathrm{N}_{m,\mathrm{z}})^{\sigma})^{\beta^{y}_{m} - \beta^{y\prime}_{m}}(\mathrm{A}_{\mathrm{z}})^{\beta^{a}_{m}} \end{aligned} \tag{9}$$

In other words, the own-gender expected earnings term, instead of capturing only the utility value of earnings per se (β^{y}), captures the negative of the value of expected leisure time ($\tilde{\beta}^{l}$) as well.

Rewriting variables in log form, (9) becomes,

$$\begin{aligned} \mathrm{u}_{w,\mathrm{z}} = \bar{\mathrm{u}}'_{w} + \beta^{w}_{w} \cdot \mathrm{y}_{w,\mathrm{z}} + \beta^{m}_{w} \cdot (\mathrm{y}_{m,\mathrm{z}} + \sigma(\mathrm{n}_{m,\mathrm{z}} - \mathrm{n}_{w,\mathrm{z}})) + \beta^{a}_{w} \cdot \mathrm{a}_{\mathrm{z}} \\ \mathrm{u}_{m,\mathrm{z}} = \bar{\mathrm{u}}'_{m} + \beta^{m}_{m} \cdot \mathrm{y}_{m,\mathrm{z}} + \beta^{w}_{m} \cdot (\mathrm{y}_{w,\mathrm{z}} + \sigma(\mathrm{n}_{w,\mathrm{z}} - \mathrm{n}_{m,\mathrm{z}})) + \beta^{a}_{m} \cdot \mathrm{a}_{\mathrm{z}} \end{aligned} \tag{10}$$

$$\begin{aligned} \beta^{w}_{w} = \beta^{y}_{w} - \tilde{\beta}^{l}_{w} \ , \ \beta^{m}_{m} = \beta^{y}_{m} - \tilde{\beta}^{l}_{m} \\ \beta^{m}_{w} = \beta^{y}_{w} - \beta^{y\prime}_{w} \ , \ \beta^{w}_{m} = \beta^{y}_{m} - \beta^{y\prime}_{m} \end{aligned} \tag{11}$$

$$\mathrm{y}_{w,\mathrm{z}} = log(\mathring{\mathrm{Y}}_{w,\mathrm{z}}) \ , \ \mathrm{y}_{m,\mathrm{z}} = log(\mathring{\mathrm{Y}}_{m,\mathrm{z}}) \ , \ \mathrm{n}_{w,\mathrm{z}} = log(\mathrm{N}_{w,\mathrm{z}}) \ , \ \mathrm{n}_{m,\mathrm{z}} = log(\mathrm{N}_{m,\mathrm{z}})$$

$$\bar{\mathrm{u}}'_{w} = log(\bar{\mathrm{U}}_{w}) \ , \ \bar{\mathrm{u}}'_{m} = log(\bar{\mathrm{U}}_{m}) \ , \ \mathrm{a}_{\mathrm{z}} = log(\mathrm{A}_{\mathrm{z}})$$

Because utility is immune to monotonic transformation, there is no need to write $log(\mathrm{u}_{\mathrm{z}})$ or $exp(\mathrm{u}_{\mathrm{z}})$ in place of u_{z} in transition from (9) to (10), or vice versa. Moreover, recall that all variables have also (unwritten) t subscripts. It would be natural to divide each earnings term $\mathring{\mathrm{Y}}$ by a deflator specific to each time period t, to make earnings real rather than nominal in comparison over time. However, like in (6), this would make no difference, because the deflator terms would be absorbed in the intercept terms, $\bar{\mathrm{u}}'$, which in the panel context are absorbed in period fixed effects.

1.2.4 Symmetry of Preferences

It would be natural to expect women and men each to value the other's earnings positively, but nonetheless each to prioritize their own earnings above all. This is the basic intuition underlying the hypothesis of Symmetric Preferences. As expressed in (10), preferences can be summarized via four main preference parameters β, which are women's and men's own and cross effects.

$$\text{women's attraction to} \left\{ \overbrace{\underbrace{\begin{matrix} \beta_w^w & \beta_m^w \\ \beta_w^m & \beta_m^m \end{matrix}}_{\text{men's earnings}}}^{\text{women's earnings}} \right\} \text{men's attraction to}$$

Roughly speaking, what I mean by Symmetry is the idea that the own effects are equal ($\beta_w^w = \beta_m^m$) and the cross effects are equal ($\beta_w^m = \beta_m^w$).[30]

The previous section unpacked what meaning should be captured in each own and cross effect, concluding in (11). That is, each own effect is the value of earnings (β^y)

[30]This is in fact a stronger condition than my official statement of the hypothesis, (13), just maybe conceptually easier, and a special case of (13).

minus the value of leisure time ($\tilde{\beta}^l$), and each cross effect is the value of earnings (β^y) minus the value of direct control over earnings ($\beta^{y'}$) that is afforded by earning directly rather than through a spouse. What is distinct between them, therefore, comes down to the difference in importance between leisure time ($\tilde{\beta}^l$) and direct control ($\beta^{y'}$). This is explicit in the mathematical difference of the own and cross effect:

$$\begin{aligned} \lambda_w^{w-m} &= \beta_w^w - \beta_w^m = (\beta_w^y - \tilde{\beta}_w^l) - (\beta_w^y - \beta_w^{y'}) = \beta_w^{y'} - \tilde{\beta}_w^l \\ \lambda_m^{m-w} &= \beta_m^m - \beta_m^w = (\beta_m^y - \tilde{\beta}_m^l) - (\beta_m^y - \beta_m^{y'}) = \beta_m^{y'} - \tilde{\beta}_m^l \end{aligned} \tag{12}$$

By Symmetric, I would mean women and men place similar importance on leisure vis-à-vis direct control. Mathematically, this Symmetry would coincide with ($\lambda_w^{w-m} = \lambda_m^{m-w}$).

When comparing preference parameters across different time periods, it is important again to remember that utility is immune to monotonic transformation. Therefore a proportional increase or decrease in the magnitudes of all preference parameters would not actually mean anything; it is only the relative values that have meaning. To enable comparison across years, I assume that the sum of all four own and cross effect β parameters has constant meaning. Dividing each particular preference parameter by this sum, therefore, quantifies preferences in a way that is comparable across different time periods. Preferences are Asymmetric in so far as ($\lambda_w^{w-m} - \lambda_m^{m-w}$) deviates from zero. To quantify the magnitude of said Asymmetry, therefore, I divide ($\lambda_w^{w-m} - \lambda_m^{m-w}$) by the aforementioned sum. That is,

$$\Lambda = \frac{\beta_w^w - \beta_w^m + \beta_m^w - \beta_m^m}{\beta_w^w + \beta_w^m + \beta_m^w + \beta_m^m} \tag{13}$$

The hypothesis of Preference Symmetry is that $\Lambda = 0$. A negative value of Λ would indicate that men's earnings are prioritized over women's, which under my model is conceptually equivalent to the idea that women value direct control over earnings ($\beta^{y'}$) vis-à-vis expected leisure time ($\tilde{\beta}^l$) relatively *less* than men do. A positive value of Λ would indicate the opposite - that women's earnings are prioritized over men's, and that women value direct control over earnings, above leisure, relatively *more* than men do.

Since cultural norms may play a role in preferences, Symmetry is less likely to hold in the earlier years of the panel. However, if asymmetry in preferences arises from asymmetry in norms, then it should diminish as asymmetry in norms diminishes. The huge increase in women's earnings opportunities over time can only have shifted some cultural attention towards women's earnings opportunities relative to men's. As I show in Table 1.2, $\bar{\mathrm{Y}}_{jw,\mathrm{t}}/\bar{\mathrm{Y}}_{jm,\mathrm{t}}$ increased by more than 100% from 1970 to the present.[31] Explicitly, also, Donnelly, Twenge and Clark (2016) find that attitudes toward working mothers versus working fathers became more egalitarian over time, in every decade from 1970 to the 2010s. If $\Lambda < 0$ earlier in time, then, Symmetry in a broader sense must predict that the magnitude of the asymmetry Λ would diminish later in time.

1.2.5 Elasticities of Net Migration

In addition to being preference parameters by definition under the model, the parameters β are also, by derivation under the model, equivalent or closely related to elasticities of net migration (with respect to expected earnings). Elasticities of net migration are what determine effects on sex ratios, which constitute the other central subject question of the paper (in addition to Symmetry or Asymmetry of preferences). To arrive at the migration elasticities, combine equations (3) and (10), substituting out $\mathrm{u_z}$:[32]

$$\begin{aligned} \mathrm{n}_{w,\mathrm{z}} - \tilde{\mathrm{b}}_{w,\mathrm{z}} &= \beta^{w}_{w} \cdot \mathrm{y}_{w,\mathrm{z}} + \beta^{m}_{w} \cdot (\mathrm{y}_{m,\mathrm{z}} + \sigma(\mathrm{n}_{m,\mathrm{z}} - \mathrm{n}_{w,\mathrm{z}})) + \beta^{a}_{w} \cdot \mathrm{a_z} + \mathrm{X}_{w} \\ \mathrm{n}_{m,\mathrm{z}} - \tilde{\mathrm{b}}_{m,\mathrm{z}} &= \beta^{m}_{m} \cdot \mathrm{y}_{m,\mathrm{z}} + \beta^{w}_{m} \cdot (\mathrm{y}_{w,\mathrm{z}} + \sigma(\mathrm{n}_{w,\mathrm{z}} - \mathrm{n}_{m,\mathrm{z}})) + \beta^{a}_{m} \cdot \mathrm{a_z} + \mathrm{X}_{m} \end{aligned} \tag{14}$$

[31] $\bar{\mathrm{Y}}_{jw,\mathrm{t}}/\bar{\mathrm{Y}}_{jm,\mathrm{t}}$ is the ratio of women's (w) national average earnings ($\bar{\mathrm{Y}}$) to men's (m) national average earnings, amongst (j) non-institutionalized heterosexuals of age 18-64.

[32] $\mathrm{X}_{w} = \bar{\mathrm{u}}_{w} - \bar{\mathrm{u}}'_{w}$ is an intercept.

$$\tilde{\mathrm{b}}_{w,\mathrm{z}} = log(\,exp(\hat{\beta}^b_w)\mathrm{b}_{w,\mathrm{z}} + 1 - \mathrm{b}_{w,\mathrm{z}})\ ,\ \tilde{\mathrm{b}}_{m,\mathrm{z}} = log(\,exp(\hat{\beta}^b_m)\mathrm{b}_{m,\mathrm{z}} + 1 - \mathrm{b}_{m,\mathrm{z}})$$

Women's elasticity of net migration with respect to men's expected earnings (for example) is $d\,\mathrm{n}_{w,\mathrm{z}}/d\,\mathrm{y}_{m,\mathrm{z}}$, because $\mathrm{n}_{w,\mathrm{z}}$ and $\mathrm{y}_{m,\mathrm{z}}$ are already in log terms. $d\,\mathrm{n}_{w,\mathrm{z}}/d\,\mathrm{y}_{m,\mathrm{z}}$ cannot be evaluated here due to the presence of n_z terms on both sides of both equations in the system (37). However, this system has a unique linear solution, which is,

$$\begin{aligned}\mathrm{n}_{w,\mathrm{z}} &= \gamma^w_w \cdot \mathrm{y}_{w,\mathrm{z}} + \gamma^m_w \cdot \mathrm{y}_{m,\mathrm{z}} \;+\; \text{[unrelated terms]}\\ \mathrm{n}_{m,\mathrm{z}} &= \gamma^m_m \cdot \mathrm{y}_{m,\mathrm{z}} + \gamma^w_m \cdot \mathrm{y}_{w,\mathrm{z}} \;+\; \text{[unrelated terms]}\end{aligned} \tag{15}$$

$$\gamma^w_w = \frac{\beta^w_w + \sigma\beta^w_m(\beta^w_w + \beta^m_w)}{1 + \sigma(\beta^w_m + \beta^m_w)} \quad , \quad \gamma^m_m = \frac{\beta^m_m + \sigma\beta^m_w(\beta^m_m + \beta^w_m)}{1 + \sigma(\beta^w_m + \beta^m_w)}$$
$$\gamma^m_w = \frac{\beta^m_w + \sigma\beta^m_w(\beta^m_m + \beta^w_m)}{1 + \sigma(\beta^w_m + \beta^m_w)} \quad , \quad \gamma^w_m = \frac{\beta^w_m + \sigma\beta^w_m(\beta^w_w + \beta^m_w)}{1 + \sigma(\beta^w_m + \beta^m_w)}$$

Each γ is an elasticity of net migration. For example, women's elasticity of net migration with men's expected earnings is $d\ \mathrm{n}_{w,\mathrm{z}}/d\ \mathrm{y}_{m,\mathrm{z}} = \gamma^m_w$. Moreover, it is apparent in (15) that in the case of $\sigma = 0$, each γ is equivalent to the corresponding β, such as, $\gamma^m_w = \beta^m_w$.

Because $\mathrm{n}_{w,\mathrm{z}} - \mathrm{n}_{m,\mathrm{z}}$ is the log sex ratio (population ratio of women to men in location z), the gender difference of the equations (15) yields elasticities of the sex ratio with respect to expected earnings. That is,

$$\mathrm{n}_{w,\mathrm{z}} - \mathrm{n}_{m,\mathrm{z}} = (\gamma^w_w - \gamma^w_m) \cdot \mathrm{y}_{w,\mathrm{z}} + (\gamma^m_w - \gamma^m_m) \cdot \mathrm{y}_{m,\mathrm{z}} \;+\; \text{[unrelated terms]} \tag{16}$$

$(\gamma^w_w - \gamma^w_m)$ and $(\gamma^m_w - \gamma^m_m)$ are the elasticities of the sex ratio with respect to women's and men's expected earnings, respectively. It follows that the difference of $(\gamma^w_w - \gamma^w_m)$ and $(\gamma^m_w - \gamma^m_m)$,

$$\gamma^{w-m}_{w-m} = \gamma^w_w - \gamma^m_w - \gamma^w_m + \gamma^m_m \tag{17}$$

is the elasticity of the sex ratio with respect to the earnings ratio. A negative value of γ^{w-m}_{w-m} would be counterintuitive, indicating that women are more attracted than men are to places in which women's earnings are below par relative to men's.[33] Of course, it would be counterintuitive primarily in a sense of preferences, rather than of migration elasticities per se. This is because γ_{w-m}^{w-m} is also equal (in the $\sigma = 0$ case) to the sum of λ_w^{w-m} and λ_m^{m-w} as defined in (12).[34] A final object of interest is the sum of $(\gamma_w^w - \gamma_m^w)$ and $(\gamma_w^m - \gamma_m^m)$,

$$\gamma_{w-m}^{w+m} = \gamma_w^w + \gamma_w^m - \gamma_m^w - \gamma_m^m \tag{18}$$

which is a way of summarizing the effect of earnings overall on the sex ratio. A positive value of γ_{w-m}^{w+m} would indicate that women are more attracted to earnings *overall* than men are, which may in fact be more important in driving sex ratio imbalance than earnings *differences* between women and men (as summarized rather by γ_{w-m}^{w-m}) are in driving sex ratio imbalance.

1.3 Identification

Several identification hurdles arise with regard to estimating (37). I summarize these hurdles in five categories: reverse causality, correlated explanatory variables, omitted explanatory variables, heterogeneity, and inertial frictions. My responses to each in turn are: Shift-Share (Bartik) instruments, maximizing the size of the dataset, regional fixed effects and year by region effects, restrictions by different demographic groups j, and serial autocorrelation.

[33]This counterintuitive result ($\gamma_{w-m}^{w-m} < 0$) is basically the finding of Edlund (2005), though in different terms.

[34]Most intuitive would be if both λ_w^{w-m} and λ_m^{m-w} were positive, meaning that women and men each prioritize their own earnings primarily, rather prioritizing the other's earnings primarily. However, it is less counterintuitive for either λ to be negative, and to any magnitude, specifically in light of the reasoning being presented in (12), that women may prioritize men's earnings (or vice versa) for the reason of enabling leisure time. Note that the sum of λ_w^{w-m} and λ_m^{m-w} does not coincide with the parameter of Asymmetry, Λ, which is determined rather by the difference of λ_w^{w-m} and λ_m^{m-w}.

1.3.1 Shift-Share (Bartik) Instruments

In (37), the response (left hand side) variables are essentially population counts, while the main explanatory (right hand side) variables of interest are expected earnings. The basic economic idea coinciding with (37) is that people go where the money is; therefore population counts become higher where expected earnings are higher. However, population counts may also have reverse effects on expected earnings. In a classic sense of downward sloping labor demand, the presence of more people may depress employment rates and/or wages, resulting in lower earnings. Meanwhile, agglomeration effects to productivity may act in exactly the opposite direction, causing average earnings to increase with population rather than decrease.[35]

To get around any such effects of population counts on expected earnings, I construct a set of Shift-Share (Bartik) instruments,[36] one for each demographic group j.[37] The instrument, for each j, interacts the lagged fractional breakdown of employment opportunities in each location z across occupational categories, with the corresponding vector of national average earnings for each category. This arrives at a measure of expected earnings that is decoupled from the effects of local population counts in each location z, depending rather only on the historical structure of the economy in each.

The component of the instrument that is specific to each location z is what I call the Occupational Profile (OP). The OP of each Commuting Zone z, in each time period t, for each demographic type j, is a vector of 14 fractions summing to one, giving the occupational breakdown of employment across each in z. The 14 fractions coincide with the 14 Occupations (occupational categories) listed in Table 1.6. That

[35]Diamond (2016) indeed recovers a positive elasticity of labor demand for college educated workers, indicating that agglomeration effects outweigh the classic downward slope of labor demand. Glaeser and Maré (2001) discuss such agglomeration effects, emphasizing information spillover as a likely mechanism by which productivity may increase with the labor quantity (population), especially of skilled workers.

[36]The name Bartik comes from to Bartik (1991), but refers to a broad class of instruments. See Bound and Holzer (2000), Diamond (2016), Notowidigdo (2020), and Shenhav (2021) for examples.

[37]j here as a notation subsumes restrictions by gender (w or m), as well as any other demographic restrictions, such as age range, employment status and marital status.

is, the OP for j,t,z is the vector,

$$\left\{\frac{\mathrm{NE}_{\mathrm{o,j,t,z}}}{\mathrm{NE}_{\mathrm{j,t,z}}}\right\}_{\mathrm{o}}$$

where $\mathrm{NE}_{\mathrm{o,j,t,z}}$ is the count of type j people employed in Occupation o, and $\mathrm{NE}_{\mathrm{j,t,z}}$ is the count of type j people employed in any Occupation. To arrive at the instrument, the OP is interacted with the vector,

$$\left\{\mathrm{Y}_{\mathrm{o,j,t}}\right\}_{\mathrm{o}}$$

where $\mathrm{Y}_{\mathrm{o,j,t}}$ is the national average post-tax earnings of type j people who identify Occupation o as their habitual occupation (even if not employed).

I assume that OPs are a first mover of the system of supply and demand, arising from a combination of exogenous geographical, historical, and technological factors. In other words, OPs do not depend on population (labor) supply responses, that is, migration choices. For example: though migrant choice influences the total population of Boston, migrant choice does not influence the probability that a migrant to Boston, if employed, will be employed in finance. In case this is considered questionable, I lag OPs by one time period. The one period lag allows OPs to update for relevance to the nature of economic activity in the current time period, while remaining clearly causally prior to migrant choice in the current period. The instrument therefore is,

$$\tilde{\mathrm{y}}_{\mathrm{j,t,z}} = log(\tilde{\mathrm{Y}}_{\mathrm{j,t,z}}) \ , \ \tilde{\mathrm{Y}}_{\mathrm{j,t,z}} = \sum_{\mathrm{o}} \frac{\mathrm{NE}_{\mathrm{o,j,t-1,z}}}{\mathrm{NE}_{\mathrm{j,t-1,z}}} \cdot \mathrm{Y}_{\mathrm{o,j,t}} \tag{19}$$

where t = (0, 1, 2, 3, 4, 5, 6) indicate survey years (1950, 1970, 1980, 1990, 2000, 2010, 2019) respectively.

In the case of $\sigma > 0$, there is an additional source of reverse causality. Repeating

(37),

$$\mathrm{n}_{w,\mathrm{z}} - \tilde{\mathrm{b}}_{w,\mathrm{z}} = \beta_w^w \cdot \mathrm{y}_{w,\mathrm{z}} + \beta_w^m \cdot (\mathrm{y}_{m,\mathrm{z}} + \sigma(\mathrm{n}_{m,\mathrm{z}} - \mathrm{n}_{w,\mathrm{z}})) + \beta_w^a \cdot \mathrm{a}_\mathrm{z} + \mathrm{X}_w$$

$$\mathrm{n}_{m,\mathrm{z}} - \tilde{\mathrm{b}}_{m,\mathrm{z}} = \beta_m^m \cdot \mathrm{y}_{m,\mathrm{z}} + \beta_m^w \cdot (\mathrm{y}_{w,\mathrm{z}} + \sigma(\mathrm{n}_{w,\mathrm{z}} - \mathrm{n}_{m,\mathrm{z}})) + \beta_m^a \cdot \mathrm{a}_\mathrm{z} + \mathrm{X}_m$$

The cross effects β_w^m and β_m^w would be confounded by reverse causality, not only in the sense that population counts $\mathrm{n}_{w,\mathrm{z}}$ and $\mathrm{n}_{m,\mathrm{z}}$ may have effects on expected earnings $\mathrm{y}_{w,\mathrm{z}}$ and $\mathrm{y}_{m,\mathrm{z}}$, but also because $\mathrm{n}_{w,\mathrm{z}}$ and $\mathrm{n}_{m,\mathrm{z}}$ are themselves present on the right hand side of the equation as well as the left.[38] However, the instruments $\tilde{\mathrm{y}}_{\mathrm{t},\mathrm{z}}$ get around both of these problems. The endogenous variable on β_w^m is in fact the whole object $(\mathrm{y}_{m,\mathrm{z}} + \sigma(\mathrm{n}_{m,\mathrm{z}} - \mathrm{n}_{w,\mathrm{z}}))$. Therefore only one instrument for the whole object is needed, namely $\tilde{\mathrm{y}}_{m,\mathrm{z}}$, to identify β_w^m. Likewise $\tilde{\mathrm{y}}_{w,\mathrm{z}}$ instruments for the whole object $(\mathrm{y}_{w,\mathrm{z}} + \sigma(\mathrm{n}_{w,\mathrm{z}} - \mathrm{n}_{m,\mathrm{z}}))$, identifying β_m^w. This recovers the *preference parameters* β, which in the $\sigma > 0$ case are distinct from the *migration elasticities* γ; the latter are then computed via the formulae in (15). The migration elasticities quantify what equilibrium population shifts would occur in response to counterfactual changes in expected earnings, as people respond both to their own preferences and to others'.

1.3.2 Correlated Explanatory Variables

It is of course the case that women's and men's expected earnings, which are the main explanatory variables of interest, are correlated with one another. That is to say, places z in which women's earnings are above average, relative to women nationally, also tend to places in which men's earning are above average relative to men nationally. This is intuitively concerning, because it seems unclear how the researcher would in any case determine whether, for example, a single woman moves to Denver for her own career prospects in Denver, or rather for her potential spouses' career prospects in Denver, both of which are good.

[38] This extreme form of endogeneity is why σ itself cannot be estimated in the context of the model, so needs rather to be calibrated.

Alone amongst the hurdles discussed in this (Identification) section, however, the fact that women's and men's expected earnings are correlated with one another actually is not a genuine identification issue, but rather only seems to be at first thought. It is almost always the case, in every econometric estimation, that explanatory variables are correlated with one another. An identification problem would arise rather if such a variable - that both belongs in the model as an explanatory variable, and is correlated with others - were *omitted*. And in such a case, the remedy would be to include said variable.

Intuitively, the independent effects of women's and men's expected earnings are identified, even though the two are correlated with another, because women's earnings are above par relative to men's in some places z, and below par in others.[39] The higher is the correlation, the less information any one place z observation will contain about their independent effects. But this is less information in the same sense that a larger dataset contains more information than a smaller dataset. In other words, it is if anything a situation of *small sample bias*, which is a universal feature of all estimations, and the only remedy to which is to favor using more data rather than less.

1.3.3 Fixed Effects

As is almost always the case in any real setting, the potential for omitted variables to yield bias is the truest and most intractable identification issue. In this setting, miscellaneous factors of interest are called *amenities*, $a_{t,z}$. Amenities are characteristics of location options z, other than expected earnings, that people on average would value in deciding where to live. Explicit amenities can be included, such as weather, crime rates, air quality, and many others. If any of these are omitted, and correlated with expected earnings, it is possible that the estimation would falsely read their effects as effects of expected earnings. However, because it is impossible to observe in

[39]Of course, it is required that they are not *perfectly* correlated.

data everything that makes some places z more desirable to live in than others, the issue could not ever be settled conclusively.

For simplicity, I rely on a standard method that is meant to (plausibly/perhaps) absorb all relevant unobserved information, namely, fixed effects. Because not all amenities are of a nature that is fixed over time, I use region-by-year effects as well regional fixed effects. For example, the weather in New York is about equally worse than the weather in Florida regardless of the time period, but the same cannot always be said of things like cultural opportunities. I use the nine regional divisions used by the Census,[40] and further split each of these into their urban and rural Commuting Zones, for a total of 18 regions.[41] I include an effect for each of these, in each of the 6 decadal time periods, for a total of 108 effects. I use this array of 108 effects, $X_{t,z}$, in place of the amenities term, $a_{t,z}$. Because $a_{t,z}$ is meant to represent the grand sum total of things of value other than expected earnings, it can of course be fairly questioned whether $X_{t,z}$ is up to the task. Most important would be amenities that are correlated *differently* with women's as opposed to men's expected earnings. Exploring such factors may be a fruitful avenue for future work.

1.3.4 Heterogeneity

The model uses population counts and expected earnings on a large scale to reveal average preferences. This naturally brings with it the possibility that the model should apply differently, or not at all, to various subsets of the population, which should therefore be distinguished from the whole. Many potential concerns are of this nature. This is the role of restricting by demographic characteristics, j, which is a constant theme of the model starting from equation (1). j specifies groups of people whom the model assumes to share common parameters of *average* utility. It is tautological that any group of people can share a common average; but averages are more meaningful when summarizing values that are similar to one another.

[40]These nine regional divisions have the benefit of being about the same size as one another.
[41]See Appendix section 1.9.1 for details.

Some demographic restrictions j should be applied across the board, because the model does not make sense in application to people not meeting them. Other restrictions j are optional, in that the model applies either way, but perhaps with very different parameters of average utility. An obvious example of the former is sexual orientation. The model does not apply at all to non-heterosexuals, because a core assumption is that expected earnings of expected spouses are captured by expected earnings of the opposite gender. Such essential restrictions - those that should be applied across the board - together form the baseline group, j = $\jmath$. Aside from sexual orientation, characteristics of $\jmath$ mainly revolve around freedom of choice. The model assumes that all addresses are choices, and therefore reflect preferences: everyone had freedom to choose to live elsewhere than they did. This is not true of institutionalized people, and people below the age of 18. I also exclude those above 64, mainly to avoid any significant heterogeneity in survivorship, because death is not a choice. The baseline group $\jmath$ therefore is the set of all non-institutionalized heterosexuals of age 18-64.

With regard to optional demographic restrictions j, it is appropriate to present multiple sets of results for comparison, that is, results for multiple different j groups. However, there is a prodigious variety of potential such optional restrictions. For practicality, is necessary to select a handful that are most important. The most important optional restriction, I believe, is marital status. Marital status is indeed optional rather than essential, which means that the model applies in fundamentally the same way, regardless of marital status. The model regards married people as well as singles as independent agents, more similar to singles. Both married people and singles weigh both their own expected earnings, and their *expected* spouses' expected earnings, in deciding where to live. For a married woman, the expected spouse is likely to be the current husband. But in so far as location has bearing on earnings, the expected earnings of the current husband, per location, are equal to the expected earnings of men (of the same j group) in each location, which are also the expected earnings of single women's expected spouses. Therefore the pool of both married people and singles can be taken as a single j group, j, with common parameters of average utility. However, it may be reasonable to expect that married people, especially women, would place more weight on their spouses' earnings prospects than singles would place on their expected spouses' earnings prospects. This would manifest in the model as, for example, a lower value of β_w^m for single women as opposed to married women (or as opposed to the pool of both, j).

The second most important optional j restriction, I believe, is employment status - particularly, whether or not one is in the labor force. The model assumes that expected earnings per location capture the bearing that people expect their location choice to have on their own and on their expected spouses' earnings, both in the immediate and more distant future. This assumption applies less, or at least not in the same way, to people who are not in the labor force, including students. However, the choice of whether or not to be in the labor force may, in many cases, be subsumed under the main choice of the model, location choice. For example, in so far as women on average prefer to work less than men (or to not work at all), women on average will tend to choose locations accordingly - those in which men's expected earnings are above par relative to women's. In such cases, there is no need to restrict by employment status, which can be viewed rather as a distributional outcome of location choices amongst people (for each gender) of common type j, and not important in distinction from expected earnings as a whole.

In summary, I believe j itself (by gender) - the set of all non-institutionalized heterosexuals of age 18-64 - is the most meaningful group for whom to estimate common parameters of average utility. Estimating for this large group also has the advantage of more fully capturing net migration elasticities, and hence effects on sex ratios. However, there are good reasons also to believe that the parameters of average utility may differ meaningfully by marital status, and by employment status. Because I think j itself is the most important group, I present results for j, alongside results for *js* (*j* restricted to singles) and *jl* (*j* restricted to those in the labor force). It would be just as well to present results for married people in particular, alongside those for singles, but this would be somewhat redundant, as married people are of course responsible for any differences between *j* and *js*.

Marital status and employment status are by no means the *only* optional j restrictions that may be of interest. I have examined, for example, restrictions by race, education level, and different age brackets.[42] Moreover, it is not necessarily the case that the j group should be the same on both sides of the equation. For example: single women, in deciding where to live, may respond not only to the expected earnings single men, but rather to the expected earnings of married men as well, because the earnings of married men may better represent what single women would expect their future spouses to earn in the future. For simplicity, I avoid any such specifications, keeping j the same on both sides of the equation. And likewise, to limit the volume of the paper's tables, I present results only for the handful of j groups that I consider most central.

1.3.5 Inertial Frictions

An important feature of the model is that it assumes all addresses to be choices, and therefore to reveal preferences. Even a person who didn't move has made a migration choice, that is, to stay in the same place. The choice to not move undeniably does contain information about preferences: it indicates that the current location continues to present sufficiently high utility, overriding the opportunity cost of foregoing all other location options. However, part of the draw of staying in one's current place, besides its objective characteristics such as expected earnings, must arise rather from inertial frictions, such as moving costs, social connections and familiarity.

I account for inertial frictions in two ways. First, as discussed at the outset of the model, I estimate preferences to live in the state in which one was born, and

[42] Results are generally similar to those for *j* or *js*; non-whites, the college educated, and younger ages tend more like *js*.

subtract out their effects in driving location choices prior to estimating the main utility function.[43] The special benefits of living in the state in which one was born, if any, would arise from social connections, familiarity, and often a lack of relocation costs. Therefore preferences to live in one's state of birth coincide with inertial frictions, and a substantial portion of such frictions should load onto them, rather than loading onto any of the main utility parameters (β_w^m or so on). To account for any remaining inertial frictions, I impose (one lag) serial autocorrelation. This represents the fact that many people simply stay in the same place as in the previous time period, due to the effect of the inertial frictions.

1.3.6 Estimation

The parameters to be estimated are the β parameters of equation set (37). As discussed in Section 1.3.3, I replace the amenities term with the array of effects $X_{t,z}$. In notation, I now also restore the j,t subscripts, which have been hidden since equation (5).

$$\begin{aligned} \tilde{n}_{jw,t,z} &= \beta_{jw,t}^{w} \cdot y_{jw,t,z} + \beta_{jw,t}^{m} \cdot (y_{jm,t,z} + \sigma(n_{jm,t,z} - n_{jw,t,z})) + X_{jw,t,z} \\ \tilde{n}_{jm,t,z} &= \beta_{jm,t}^{m} \cdot y_{jm,t,z} + \beta_{jm,t}^{w} \cdot (y_{jw,t,z} + \sigma(n_{jw,t,z} - n_{jm,t,z})) + X_{jm,t,z} \end{aligned} \tag{20}$$

$$\begin{aligned} \tilde{n}_{jw,t,z} &= n_{jw,t,z} - \mathit{log}(\mathit{exp}(\hat{\beta}_{jw,t}^{b})b_{jw,t,z} + 1 - b_{jw,t,z}) \\ \tilde{n}_{jm,t,z} &= n_{jm,t,z} - \mathit{log}(\mathit{exp}(\hat{\beta}_{jm,t}^{b})b_{jm,t,z} + 1 - b_{jm,t,z}) \end{aligned}$$

Notice that the parameters have t subscripts. This allows that the preference parameters *may* change over time; whether they *do* change over time is a main question of the paper. For conciseness, I assume that the parameters are not distinct by *every* time period. Rather, each β_t has a constant value, β, for {1970, 1980, 1990}, and another constant value, $\beta + \Delta\beta$, for {2000, 2010, 2019}. Whether there is a significant difference between them - that is, whether $\Delta\beta$ is statistically significant - determines

[43] See equations (3) and (37).

whether there has been change over time. I use the symbol $\S_t$ as a binary variable indicating whether the year (t) is in the later section of the panel. $\S_t = 0$ for years {1970, 1980, 1990}, and $\S_t = 1$ for years {2000, 2010, 2019}. The equations to be estimated therefore are,

$$\begin{aligned}
\tilde{n}_{jw,t,z} &= \beta^{w}_{jw} \cdot y_{jw,t,z} + \beta^{m}_{jw} \cdot (y_{jm,t,z} + \sigma(n_{jm,t,z} - n_{jw,t,z})) \\
&\quad + \Delta\beta^{w}_{jw} \cdot y_{jw,t,z} \cdot \S_t + \Delta\beta^{m}_{jw} \cdot (y_{jm,t,z} + \sigma(n_{jm,t,z} - n_{jw,t,z})) \cdot \S_t + X_{jw,t,z} \\
\tilde{n}_{jm,t,z} &= \beta^{m}_{jm} \cdot y_{jm,t,z} + \beta^{w}_{jm} \cdot (y_{jw,t,z} + \sigma(n_{jw,t,z} - n_{jm,t,z})) \\
&\quad + \Delta\beta^{m}_{jm} \cdot y_{jm,t,z} \cdot \S_t + \Delta\beta^{w}_{jm} \cdot (y_{jw,t,z} + \sigma(n_{jw,t,z} - n_{jm,t,z})) \cdot \S_t + X_{jm,t,z}
\end{aligned} \tag{21}$$

I estimate (21) via Feasible Generalized Least Squares (FGLS) with instrumental variables. FGLS imposes one lag autocorrelation, as discussed in Section 1.3.5, both estimating the autocorrelation coefficient and correcting for it. As discussed in Section 1.3.1, $\tilde{y}_{jw,t,z}$ instruments for $y_{jw,t,z}$ and $(y_{jw,t,z} + \sigma(n_{jw,t,z} - n_{jm,t,z}))$, while $\tilde{y}_{jm,t,z}$ instruments for $y_{jm,t,z}$ and $(y_{jm,t,z} + \sigma(n_{jm,t,z} - n_{jw,t,z}))$.

1.4 Data

A dataset appropriate to answer the question at hand is necessarily large in scope. Because women's and men's expected earnings are correlated,[44] large cross sections are needed to estimate their independent effects. Moreover, each observation in the cross section is a geographical cell - and therefore a large number of individual person observations are needed, separately within every cell, in order to accurately calculate expected earnings and other measures within each. This is a challenge because, in most datasets, geographical location is redacted up to a high level of aggregation (such as state, rather than county). And few datasets would have enough observations to occupy a large number of geographical cells, with a large number of observations

[44] In other words, Commuting Zones (CZs) in which women's earnings are relatively high (compared to women's earnings in other CZs) also tend to be places in which men's earnings are relatively high (compared to men's earnings in other CZs).

within each.[45] Finally, in order to investigate change over time, the dataset must do all of the above for multiple time periods, and in reference to geographical cells that are fixed over time.

IPUMS USA, sourced from the Decennial Census and American Community Survey (ACS), is the largest publicly available compilation of microsamples (samples in which the observations are of individual persons). IPUMS provides 1% samples of the entire US population. That is, the total number of observations in the sample per time period is about 1% of the national population, so about 3 million (or 2 million, for earlier periods). Observations of individual persons are necessary for flexibility in variable creation. For example, a dataset of county level averages (rather than of individuals) may have the average earnings of women in each county, and the average earnings of singles, but not the average earnings of single women. However, I also use Census and ACS summaries at the county level, in addition to IPUMS. In general, IPUMS provides more specific information (variables), while the county level summaries provide more complete coverage. I combine them to yield a panel dataset with a combination of specificity and coverage that neither has on its own. See Appendix sections 1.9.1 and 1.9.2 for details.

To constitute the location choice options (z) referenced in the model, it is necessary to use geographical units that meaningfully capture local labor and marriage markets. Although Metropolitan Statistical Areas (MSAs) are the most traditional geographical unit for this purpose, Commuting Zones (CZs) have three major advantages.[46] First, CZs are more comparable over time from 1950 to the present. This is because CZs are defined as groups of counties, and county boundaries have remained mostly fixed, while MSA boundaries have been repeatedly adjusted. Third, CZs cover the entire United States, no matter how remote, while MSAs cover only urban areas. Third, CZs are delineated based on actual commuting patterns (albeit in 1990),[47]

[45]See Molloy, Smith and Wozniak (2011) for a discussion of the various datasets that can be used to study migration in the US.

[46]See Dorn (2009) for more discussion.

[47]See Tolbert and Sizer (1996).

which means they are ideal for representing actual labor and marriage markets. I use the set of 722 CZs that are coterminous with the entire continental US (48 states + DC). Each CZ is 4-5 counties on average, so the 722 CZs cover about 3,200 counties in total.

1.5 Descriptive Statistics

I present descriptive facts primarily to help familiarize the reader with the data. Table 1.1 concretely lists some Commuting Zones (CZs), including the highest and lowest in population and average earnings. The 722 CZs cover the entire mainland United States, no matter how remote. Each CZ is named after a city or town at its epicenter, but extends beyond the epicenter to a large area around it. For example, the Boston CZ contains most of eastern Massachusetts. The New York City CZ includes all of Long Island, plus a few counties to the north. The entirety of New Jersey comprises two CZs, Newark (which I rename Northern NJ), and Brick Township (which I rename Southern NJ). CZs are usually contained within state boundaries, but not always. The Washington DC CZ, for example, contains counties in Virginia and Maryland. Each CZ is delineated by Tolbert and Sizer (1996) based on commuting patterns in 1990.

Table 1.2 shows the national aggregate values for each variable, by year. From 1970 to 2019, the total population of the US mainland states rose from 203 million to 325 million, while the population of non-institutionalized heterosexuals of age 18-64 (j) rose from 90 million to 111 million. The national ratio of college educated women to college educated men $\mathrm{N}_{jwc,\mathrm{t}}/\mathrm{N}_{jmc,\mathrm{t}}$ is.[48] Likewise, the national ratio of women's earnings to men's, $\bar{\mathrm{Y}}_{jw,\mathrm{t}}/\bar{\mathrm{Y}}_{jm,\mathrm{t}}$, rose monotonically, from 24% in 1970, to 62% in 2019. This means that the average woman earns 62% as much as the average man in the

[48]This ratio is the average over the whole age range of 18-64, so is not representative of contemporary young adults. It is nonetheless relevant for understanding the overall state of earnings opportunities in each time period. Goldin, Katz and Kuziemko (2006) find that the college sex ratio for contemporary young adult cohorts troughed in 1947, having been more balanced in earlier decades. This may explain why the ratio for the 18-64 age range is at a low point around 1970.

most basic terms, that is, not filtered by employment status or hours of work in any way.

It is important that the relationship between women's and men's expected earnings varies across Commuting Zones: Women's earnings must be above par relative to men's in some places, and below par in others. To show this, Tables 1.3 and 1.4 summarize the distribution of the ratio of women's expected earnings to men's. Table 1.3 shows the ratio of actual earnings, while 1.4 shows the ratio of Occupational Profile (OP) earnings. In 1970, the median actual earnings gender ratio was 37.8% (in the median CZ, women's expected earnings were 37.8% of men's), with an IQR of 5.2%. In 2019, it was 69.3%, with an IQR of 7.0%. OP earnings are more tightly distributed than actual earnings, in part because only 14 categories cannot be very specific. In 1970, the median OP earnings gender ratio was 45.6%, with an IQR of 1.8%. In 2019, it was 70.6%, with an IQR of 3.1%. Though modest, this amount of variation should be sufficient to yield detectable effects in a large dataset.

National average earnings by occupational categories are essential components of Occupational Profile (OP) earnings, the exogenous component of expected earnings, as defined in equation (38). I calculate national averages for women and men separately, capturing the overall state of opportunities within each occupation for each gender. Opportunities may be due to the interaction of skills with technology, cultural factors such as discrimination, or both. These averages are given in Table 1.5. Like the earnings in Table 1.2, they are not filtered by employment status. They are a combined measure of the employment rate, and the earnings of the employed. Men's earnings have always been higher than women's in every category, and remain so. However, women's earnings have increased dramatically in some categories, including the most skilled categories. That is, men's work opportunities have always exceeded women's, but the difference has decreased over time.

The other components of OP earnings are the fractional breakdown of employment opportunities in each CZ, for each gender. Table 1.6 shows the national average percentage of each gender's employed population who fall into each occupational category. These essentially weight the importance of each category in determining each gender's OP earnings. For example, changes in demand for Administrative Support will have a much larger impact on women's expected earnings than Precision Production will - while for men, the opposite is true. Autor, Dorn and Hanson (2019) focus on a similar gender difference in the Manufacturing industrial category, rather than these 14 occupational categories. Industrial categories are independent of occupational categories, but do not map as well overall to skills or therefore to expected earnings.[49]

Women's concentration in the top two, highest skilled categories increased over time, likely coinciding with women's advances in education. However, women's concentration in Professional Specialty occupations always exceeded men's, even in 1950 and 1970. This does not mean that there were more women than men employed in Professional Specialty occupations, but merely that Professional Specialty occupations constituted a larger portion of women's career opportunities than it did of men's. Women's employment rates (not shown) have always been lower than men's. But conditional on working, the breakdown into various occupations is largely similar as it was in the past.[50]

1.6 Results

By estimating four preference parameters, I answer multiple questions simultaneously. As introduced in Section 1.2.3,

[49] Autor (2019) groups these same occupational categories into 3 bins - high, medium, and low skill - and 3 sub-bins within each.

[50] This may reflect biological skill endowments - men have advantages in physically intensive work, so women have corresponding comparative advantage in information intensive work.

$$\text{women's attraction to}\left\{\begin{matrix} \overbrace{\beta^w_w \quad \beta^w_m}^{\text{women's earnings}} \\ \underbrace{\beta^m_w \quad \beta^m_m}_{\text{men's earnings}} \end{matrix}\right\}\text{men's attraction to}$$

the four are preference parameters by definition - that is, they are parameters of the utility function. By derivation under the model, they also coincide with elasticities of net migration, via the formulae given in equation (15). Estimated values of the four βs therefore both reveal preferences, and predict migration behavior with respect to counterfactual changes in expected earnings across locations. Estimating the equations (21) in particular, I recover two values for each β parameter - one for the time period 1970-1990, and one for the time period 2000-2019. In addition to measuring preferences and migration elasticities, I assess the extent each changed over time.

A set of three formulae neatly summarizes how the estimated values of the four β parameters provide answers to the paper's main questions of interest about preferences and migration elasticities. As shown in Section 1.2.4, the magnitude of asymmetry in preferences between women and men essentially coincides with,

$$\beta^w_w - \beta^m_w + \beta^w_m - \beta^m_m$$

A zero value of this expression would indicate symmetry, while a negative value would indicate that men's expected earnings are prioritized over women's. As shown in Section 1.2.5, the elasticity of the women/men (population) sex ratio with respect to the women/men earnings ratio is roughly,

$$\beta^w_w - \beta^m_w - \beta^w_m + \beta^m_m$$

while the elasticity of the women/men sex ratio with respect to earnings overall is roughly,

$$\beta_w^w + \beta_w^m - \beta_m^w - \beta_m^m$$

The following Main Results section discusses the estimated values of the β parameters, and corresponding implied values of the above three expressions, both for 1970-1990 and for 2000-2019. Finally, Section 1.6.2 simulates a counterfactual tax policy change, and projects its implied impact on equilibrium sex ratio imbalance according to the 2000-2019 parameter values.

1.6.1 Main Results

I estimate equations (21), that is,

$$\begin{aligned}
\tilde{n}_{jw,t,z} &= \beta_{jw}^w \cdot y_{jw,t,z} + \beta_{jw}^m \cdot (y_{jm,t,z} + \sigma(n_{jm,t,z} - n_{jw,t,z})) \\
&\quad + \Delta\beta_{jw}^w \cdot y_{jw,t,z} \cdot \S_t + \Delta\beta_{jw}^m \cdot (y_{jm,t,z} + \sigma(n_{jm,t,z} - n_{jw,t,z})) \cdot \S_t + X_{jw,t,z} \\
\tilde{n}_{jm,t,z} &= \beta_{jm}^m \cdot y_{jm,t,z} + \beta_{jm}^w \cdot (y_{jw,t,z} + \sigma(n_{jw,t,z} - n_{jm,t,z})) \\
&\quad + \Delta\beta_{jm}^m \cdot y_{jm,t,z} \cdot \S_t + \Delta\beta_{jm}^w \cdot (y_{jw,t,z} + \sigma(n_{jw,t,z} - n_{jm,t,z})) \cdot \S_t + X_{jm,t,z}
\end{aligned}$$

where $\tilde{n}$ is log population count, minus the effect of preferences to to live in one's state of birth, w restricts to women, m restricts to men, j makes additional demographic restrictions, t indicates the time period, z indicates the particular location (Commuting Zone) option, y is log average earnings, adjusted by federal tax and local cost of living, σ is a calibrated parameter representing the effect of the population sex ratio on expected access to the opposite gender's earnings, § is a binary variable indicating whether the year is in {2000, 2010, 2019} as opposed to {1970, 1980, 1990}, and X are region by time period fixed effects. As explained in Section 1.3.6, I estimate the equations via Feasible Generalized Least Squares with one lag autocorrelation, Bartik instruments $\tilde{y}_{jw,t,z}$ and $\tilde{y}_{jm,t,z}$ (defined in Section 1.3.1), and observation weights proportional to population.

I present results for three different j groups, and two different values of σ, for a total of six specifications. The baseline j group, j = j, is non-institutionalized heterosexuals of age 18-64. As discussed in Section 1.3.4, this is the primary group of interest, because it is the set of all those who satisfy the core assumptions of the model. I also present results for js (j restricted to never married singles) and jl (j restricted to those in the labor force). The results for j may be thought of as broader average preferences over the whole group, while those for js and jl may be thought of as more specific averages for their respective subgroups. As discussed in Section 1.2.2, I am agnostic as to whether the average person would take the sex ratio itself as a choice factor in deciding where to live. Therefore I estimate assuming $\sigma = 0$, which means they do not, and also estimate assuming a positive value of σ, which means that people seek to be the scarcer sex. For this positive value I take $\sigma = 0.5$, which is on the high end of estimates of the elasticity of women's marriage rate with respect to the sex ratio.[51]

Women's preference parameter estimates are given in Table 1.7. The first row gives women's own effect (women's attraction to women's expected earnings) in 1970-1990, under each of the six specifications (columns). The second row in italics gives the change over time in women's own effect, that is, between the 1970-1990 period and the 2000-2019 period. Thus, the value of women's own effect in 2000-2019 is given by the sum of the first two rows. The third row gives women's cross effect (women's attraction to men's earnings) in 1970-1990, and the fourth row in italics gives the change over time in women's cross effect. Likewise, the first and third rows of Table 1.8 give men's own and cross effects (in that order) in 1970-1990, and the italic row under each gives the change over time in each.

The results under all six specifications follow a similar pattern. In all cases, women's cross effect (β_w^m) is the highest, followed by men's own effect (β_m^m), then women's own effect (β_w^w), and finally men's cross effect (β_m^w). The implications of this ordering, both for preferences and for elasticities of net migration, are summarized

[51] See Appendix section 1.9.4 for details.

in Table 1.9. There, Λ quantifies asymmetry of preferences; γ_{w-m}^{w-m} is the migration response of the women/men population ratio to the women/men earnings ratio; and γ_{w-m}^{w+m} summarizes the migration response of the women/men population ratio (the sex ratio) to overall (gender neutral) earnings.

Concerning asymmetry of preferences, women's cross effect is particularly striking. Not merely do women prioritize their own earnings *relatively* less than men prioritize men's earnings ($\beta_w^w - \beta_w^m + \beta_m^w - \beta_m^m < 0$). But women prioritize their own earnings less even in absolute terms, that is, $\beta_w^w < \beta_w^m$. This implies even stronger asymmetry. As mentioned earlier in this section and explained in Section 1.2.4, the summary parameter,

$$\Lambda = \frac{\beta_w^w - \beta_w^m + \beta_m^w - \beta_m^m}{\beta_w^w + \beta_w^m + \beta_m^w + \beta_m^m}$$

quantifies asymmetry of preferences in a way that can be compared across different time periods. Large values of β_w^m drive strongly negative values of Λ for both 1970-1990 and 2000-2019, indicating that men's earnings are prioritized over women's. Due to decreases in β_w^m and β_m^m, the magnitude of Λ decreased somewhat over time, but not by as much as the gender earnings gap decreased. The gender earnings ratio, $\bar{\mathrm{Y}}_{jw,\mathrm{t}}/\bar{\mathrm{Y}}_{jm,\mathrm{t}}$, is summarized in Table 1.2. Because β parameters are estimated for 1970-1990 and 2000-2019, it is appropriate to compare them to the 1980 and 2010 values of $\bar{\mathrm{Y}}_{jw,\mathrm{t}}/\bar{\mathrm{Y}}_{jm,\mathrm{t}}$. These values, 0.37 and 0.62, imply a roughly 40% decrease in the gender earnings gap. Contemporaneously, preference asymmetry Λ eroded by just under 15%.[52] Though substantial, this implies that some aspect of whatever underlies the gender asymmetry in preferences is relatively more enduring than the gender earnings gap is.

Because both of the parameters with respect to men's expected earnings (β_w^m and β_m^m) dominate those with respect to women's expected earnings (β_w^w and β_m^w), implications for sex ratio imbalance are actually of a far less extreme nature than the

[52]The erosion of Λ for never married singles (*js*) was much more extreme. However, the erosion of the gender earnings gap was much more extreme for *js* than for *j*, also.

asymmetry in preferences. That is, women go primarily where men's expected are highest; and men also go primarily where men's expected earnings are highest. Therefore sex ratios ultimately almost balance out, no matter how expected earnings are distributed across locations. However, this is not to say the effects are zero. Because women are even more attracted to men's earnings than men are to men's own earnings ($\beta_w^m > \beta_m^m$), and men are not as attracted to women's earnings as women are to women's earnings ($\beta_m^w < \beta_w^w$), effects of earnings *differences* between women and men on sex ratios (γ_{w-m}^{w-m}) are very small. However, as women seek both higher earning spouses, and higher earnings for themselves, women gravitate at a substantially steeper rate than men to locations in which *overall* (gender neutral) expected earnings are highest. Estimates of (γ_{w-m}^{w+m}) in Table 1.9 range from 0.04 to 0.33. A γ_{w-m}^{w+m} value of 0.10 would mean that a 10% increase in expected earnings, in any given location, would induce net migration that would result in a 1% increase in the ratio of women to men.[53]

1.6.2 Counterfactual

The above results pertaining to elasticities of net migration, particularly γ_{w-m}^{w+m}, suggest that policies that would reduce (gender neutral) earnings inequality across locations, such as between urban and rural areas, would have dampening effects on sex ratio imbalance. To see this in action, I simulate a counterfactual Universal Basic Income (UBI) redistribution of 5% of aggregate earnings. Even a flat tax funded UBI would effect a progressive redistribution. That is, it would shift some earnings away from higher earning people (locations) and towards lower earning people (locations). In the flat tax case, the UBI would collect $\tau \cdot \mathrm{Y_i}$ from each individual i, and return an equal share of the total collected, $(\sum_{\mathrm{i}} \tau \cdot \mathrm{Y_i})/\mathrm{N} = \tau \cdot \bar{\mathrm{Y}}$. Therefore the net gain to each individual is $\tau \cdot (\bar{\mathrm{Y}} - \mathrm{Y_i})$, meaning everyone with below average earnings gains, and everyone with above average earnings loses.

[53] In reference to singles (js) particularly, this effect may be about five times as large.

Rather than a flat tax, I assume that the UBI would be funded by a tax schedule that is itself progressive. Particularly, I use the tax function of Heathcote, Storesletten and Violante (2017), which is a good match to the actual tax schedule used in the US.[54] Of course, this makes the counterfactual UBI program even more progressive than if it were funded by a flat tax. By taking from the rich and giving to the poor, the UBI flattens (reduces inequality in) the distribution of earnings across individuals. This in turn flattens the distribution of earnings averages across locations, as some locations are characterized by a preponderance of high earning individuals, and others by a preponderance of low earning individuals.

Prior to estimating the model, I collected 20% of the aggregate pre-tax earnings in the raw data, using the Heathcote, Storesletten and Violante (2017) function, and redistributed half of that (10% of the aggregate) equally as de facto in kind benefits. For the counterfactual simulation, I shift this 10% up or down by 5%. In Table 1.10, the middle (10%) column shows the actual sex ratio imbalance in the data.[55] The left and right columns show counterfactual sex ratio imbalance. To arrive at the left (5%) column, I collect 15% of aggregate earnings as tax rather than 20%, and distribute 5% equally rather than 10%. I then calculate, according to the β parameter estimates from the model estimation, implied net migration to and from each Commuting Zone. Likewise, to arrive at the right (15%) column, I collect 25% of aggregate earnings as tax, and distribute 15%.

The counterfactual results show that the 5% redistribution of earnings, in either direction, would coincide with about a 1% change in sex ratio imbalance in the median Commuting Zone. Specifically, the 5% progressive redistribution would decrease sex ratio imbalance by about 1%, and the regressive redistribution would increase sex ratio imbalance by about the same amount. The summary parameter γ^{w+m}_{w-m} from Table 1.9, discussed earlier in this section, provides useful explanation for why this coun-

[54] See Appendix section 1.9.6 for details.

[55] To quantify sex ratio imbalance in each Commuting Zone z, I divide the number of excess of either gender, $|\mathrm{N}_{w,\mathrm{z}} - \mathrm{N}_{m,\mathrm{z}}|$, by the total, $(\mathrm{N}_{w,\mathrm{z}} + \mathrm{N}_{m,\mathrm{z}})$.

terfactual result occurs. Women gravitate at a steeper rate than men do away from lower earning locations, and towards higher earning locations. Therefore the UBI, by reducing differences in expected earnings across locations, reduces the frequency at which women choose particular (high earning) locations in excess of men. UBI is merely one, relatively simple, hypothetical policy. However, it represents in this capacity any policy that would reduce (gender neutral) inequality across locations.[56]

1.7 Conclusion

This paper finds several novel results. Though it has long been known that people gravitate to locations in which overall expected earnings are higher, the role of gender in migration choices has been largely unexplored. This paper shows for the first time, in a way that includes singles, that men's expected earnings are the primary driver of location choice for both women and men. Although women's expected earnings are valued positively, men's are prioritized by at least twice as much. Under the model I articulate, this indicates a gender asymmetry in preferences, with women on average preferring to earn through a spouse rather than directly, and men preferring to earn directly. Between 1970 and 2019, I find that the magnitude of this asymmetry in preferences eroded by about 15%, contemporaneously as the national gender earnings gap eroded by about 40%.

It is important for two reasons to understand gender differences in migration priorities. First, this understanding contributes to our overall bank of knowledge about gender, helping to inform theory and policy related to gender differences, such as wage gaps. Migration choice reveals preferences, conditional on the set of opportunities available to women and men in a given time period, by showing which of those available opportunities people tend to choose. Using a new methodology, I find results that agree with past experimental results, vignette experiments, and theory, corrob-

[56]The goal of reducing sex ratio imbalance does not by itself justify any policy intervention - it just may be worth considering as a collateral effect.

orating a common finding of gender asymmetry in preferences. Unlike most studies, my approach enables investigation of change over a long period of time, showing that a portion of the preferences in question may be culturally malleable.

Second, it is important to understand gender differences in migration priorities for the reason that migration can yield unbalanced sex ratios, which can adversely affect rates of crime, marriage, and single parenting. Regarding sex ratios, I find ironic results. Though both women and men prioritize men's expected earnings over women's, they do so by about the same amount. As a result, gender differences in earnings play little if any role in driving sex ratio imbalance in 2000-2019. However, women do place more priority weight on overall (men's and women's combined) expected earnings than men do. Therefore, overall earnings inequality across geographical areas does drive sex ratio imbalance, as women gravitate to locations with higher overall expected earnings more steeply than men do. More balanced sex ratios may follow as a beneficial side effect of policies that reduce overall earnings inequality, such as between urban and rural areas.

1.8 Tables and Figures

Table 1.1: Highest and Lowest Commuting Zones in 2010 Values

	$N_{j,t,z}$	CZ	State	$Y_{j,t,z}$	CZ	State
max	11,161,380	Los Angeles	CA	$37,779	Washington	DC
2	7,628,303	New York City	NY	$33,194	Northern NJ	NJ
3	5,307,103	Chicago	IL	$33,058	Boston	MA
4	3,787,730	Northern NJ	NJ	$32,926	Bridgeport	CT
5	3,631,180	Philadelphia	PA	$32,875	Southern NJ	NJ
6	3,491,887	Houston	TX	$32,854	San Francisco	CA
7	3,411,336	Washington	DC	$32,302	San Jose	CA
8	3,244,614	Boston	MA	$31,958	Minneapolis	MN
9	3,235,496	Detroit	MI	$31,872	Baltimore	MD
10	3,078,782	San Francisco	CA	$31,547	Nantucket	MA
11	2,938,426	Atlanta	GA	$31,006	Fredericksburg	VA
12	2,699,709	Seattle	WA	$30,667	Manchester	NH
13	2,588,475	Dallas	TX	$30,551	New York City	NY
14	2,541,511	Phoenix	AZ	$30,500	Denver	CO
15	2,217,851	Bridgeport	CT	$30,131	Vineyard Haven	MA
16	2,143,655	Miami	FL	$29,948	Seattle	WA
17	2,000,030	Minneapolis	MN	$29,664	Philadelphia	PA
18	1,949,606	San Diego	CA	$29,522	Rochester	MN
19	1,869,629	Sacramento	CA	$29,406	Chicago	IL
20	1,769,552	Denver	CO	$29,220	Wilmington	DE
70	572,607	Baton Rouge	LA	$26,318	Racine	WI
71	562,137	Albuquerque	NM	$26,312	Salt Lake City	UT
p90	561,667	Scranton	PA	$26,312	Springfield	IL
73	555,845	Omaha	NE	$26,281	Charlottesville	VA
74	541,660	Cape Coral	FL	$26,270	Vernal	UT
179	204,003	Eau Claire	WI	$24,106	Kalamazoo	MI
180	203,213	Charlottesville	VA	$24,102	Keene	NH
p75	198,891	Fredericksburg	VA	$24,075	Findlay	OH
182	197,856	Lubbock	TX	$24,068	Belcourt	ND
183	193,794	Mansfield	OH	$24,044	Sidney	MT
718	1,091	Coldwater	KS	$16,305	West Plains	MO
719	1,085	Scobey	MT	$16,302	Mount Pleasant	MI
720	730	Matador	TX	$16,094	Greenville	MS
721	712	Jordan	MT	$15,604	Middlesborough	KY
min	683	Murdo	SD	$15,589	Greenwood	MS

N_z is population count and Y_z is average post-tax earnings in each Commuting Zone, z. j restricts to age 18-64 non-institutionalized heterosexuals, t by year (all 2010 here).

Table 1.2: National Aggregate Values

Unit of Expression	Variable	1970	1980	1990	2000	2010	2019
10^6	$\mathrm{N_t}$	203	227	249	286	306	325
10^6	$\mathrm{N}_{j,\mathrm{t}}$	111	134	150	171	188	194
1	$\mathrm{N}_{jw,\mathrm{t}}/\mathrm{N}_{jm,\mathrm{t}}$	1.07	1.05	1.04	1.03	1.04	1.03
1	$\mathrm{N}_{jwc,\mathrm{t}}/\mathrm{N}_{jmc,\mathrm{t}}$	0.43	0.57	0.85	0.99	1.13	1.19
1	$\bar{\mathrm{Y}}_{jw,\mathrm{t}}/\bar{\mathrm{Y}}_{jm,\mathrm{t}}$	0.24	0.37	0.48	0.54	0.62	0.62
$\$10^3$ year^{-1}	$\bar{\mathrm{Y}}_{j,\mathrm{t}}$	4.42	8.63	16.7	25.4	29.9	40.5

N is national population count, $\bar{\mathrm{Y}}$ is national average pre-tax earnings. j restricts to age 18-64 non-institutionalized heterosexuals, c restricts to college educated. $\bar{\mathrm{Y}}_{jw,\mathrm{t}}/\bar{\mathrm{Y}}_{jm,\mathrm{t}}$ is the national gender earnings ratio.

Table 1.3: Distribution of Gender Earnings Ratio

	1950	1970	1980	1990	2000	2010	2019
min	17.0	24.2	29.6	39.5	46.0	47.1	46.2
p10	26.1	33.4	38.1	48.5	56.4	62.0	60.9
p25	29.0	34.8	41.1	51.8	59.4	66.4	65.7
p50	32.7	37.8	45.0	54.8	62.3	69.9	69.3
p75	36.6	40.0	48.0	58.0	65.1	73.3	72.7
p90	40.2	43.2	51.7	60.1	68.4	76.6	75.9
max	52.0	49.9	64.6	66.1	79.9	86.7	87.2

Distribution over CZs, z, of Gender Earnings Ratio, $\mathrm{Y}_{jw,\mathrm{t,z}}/\mathrm{Y}_{jm,\mathrm{t,z}}$, in 10^{-2} (%). $\mathrm{Y_z}$ is average post-tax wage and salary earnings in each Commuting Zone, z. j restricts to age 18-64 non-institutionalized heterosexuals, w to women, m to men, t by year.

Table 1.4: Distribution of Bartik Gender Earnings Ratio

	1950	1970	1980	1990	2000	2010	2019
min	–	34.4	42.6	49.3	57.9	64.8	63.2
p10	–	37.6	44.0	53.9	60.6	67.9	67.8
p25	–	39.6	44.5	54.7	61.5	68.9	69.0
p50	–	41.7	45.6	55.7	62.8	70.4	70.6
p75	–	44.3	46.3	57.5	64.3	72.3	72.1
p90	–	46.9	48.3	59.2	65.8	73.8	73.9
max	–	52.0	51.8	63.4	69.2	78.7	79.2

Distribution over CZs z of Bartik Gender Earnings Ratio, $\tilde{\mathrm{Y}}_{jw,\mathrm{t,z}}/\tilde{\mathrm{Y}}_{jm,\mathrm{t,z}}$, in 10^{-2} (%). $\tilde{\mathrm{Y}}_{\mathrm{z}}$ is Bartik earnings, as defined in (38). j restricts to age 18-64 non-institutionalized heterosexuals, w to women, m to men, t by year.

Table 1.5: Occupational Average Earnings

	1950	1970	1980	1990	2000	2010	2019
Executive; Managerial;	*0.50*	*0.76*	*0.89*	*1.06*	*1.18*	*1.31*	*1.36*
Administrative	1.60	2.06	2.01	1.97	2.02	1.98	1.92
Professional Specialty	*0.66*	*0.79*	*0.84*	*0.98*	*0.97*	*1.09*	*1.05*
	1.87	1.92	1.73	1.77	1.66	1.71	1.61
Technicians;	*0.51*	*0.71*	*0.78*	*0.92*	*0.97*	*1.06*	*1.03*
Related Support	2.04	1.77	1.57	1.47	1.57	1.60	1.70
Sales	*0.45*	*0.38*	*0.46*	*0.59*	*0.61*	*0.60*	*0.64*
	1.86	1.65	1.44	1.40	1.31	1.20	1.20
Administrative	*0.71*	*0.59*	*0.67*	*0.71*	*0.73*	*0.75*	*0.72*
Support; Clerical	1.92	1.49	1.37	1.14	1.02	0.92	0.89
Private Household	*0.40*	*0.34*	*0.36*	*0.38*	*0.35*	*0.34*	*0.36*
	1.19	0.89	0.76	0.70	0.60	0.56	0.58
Protective Service	*0.59*	*0.56*	*0.54*	*0.75*	*0.82*	*0.83*	*0.79*
	1.96	1.65	1.39	1.30	1.26	1.26	1.20
Service Except	*0.33*	*0.36*	*0.41*	*0.43*	*0.46*	*0.46*	*0.47*
Protective & Household	1.25	0.96	0.81	0.71	0.68	0.62	0.61
Farming; Fishing;	*0.08*	*0.21*	*0.26*	*0.31*	*0.35*	*0.40*	*0.41*
Forestry	0.42	0.57	0.55	0.56	0.59	0.56	0.60
Precision Production;	*0.53*	*0.63*	*0.67*	*0.71*	*0.72*	*0.73*	*0.69*
Craft; Repair	1.86	1.56	1.40	1.18	1.07	0.98	1.02
Machine Operators;	*0.56*	*0.52*	*0.60*	*0.61*	*0.62*	*0.58*	*0.63*
Assemblers; Inspectors	1.76	1.40	1.32	1.11	1.00	0.88	0.92
Transportation;	*0.49*	*0.52*	*0.52*	*0.59*	*0.65*	*0.64*	*0.58*
Material Moving	1.78	1.41	1.34	1.10	1.02	0.92	0.93
Handlers; Equipment	*0.45*	*0.48*	*0.53*	*0.51*	*0.52*	*0.47*	*0.50*
Cleaners; Laborers	1.36	1.01	0.95	0.80	0.74	0.64	0.68
Military	*0.56*	*0.68*	*0.65*	*0.81*	*0.86*	*0.96*	*1.02*
	1.21	1.04	1.08	1.08	1.06	1.23	1.11

Occupational averages are scaled by grand national average: *Women's* ($\mathrm{Y}_{\mathrm{o},jw,\mathrm{t}}/\mathrm{Y}_{j,\mathrm{t}}$) italic, Men's ($\mathrm{Y}_{\mathrm{o},jm,\mathrm{t}}/\mathrm{Y}_{j,\mathrm{t}}$) non-italic. Y_{o} is average post-tax wage and salary earnings (inclusive of non-employed), in each occupational category, o, as listed. j restricts to age 18-64 non-institutionalized heterosexuals, w to women, m to men, t by year. See equation (38).

Table 1.6: Occupational Shares of the Employed National Population

	1950	1970	1980	1990	2000	2010	2019
Executive; Managerial;	*4.75*	*6.44*	*8.39*	*12.6*	*12.7*	*13.7*	*14.8*
Administrative	11.6	14.0	13.6	13.7	14.0	14.7	14.8
Professional Specialty	*11.7*	*14.0*	*15.6*	*18.0*	*21.2*	*22.5*	*23.6*
	5.56	9.42	9.42	12.4	13.6	13.8	14.1
Technicians;	*1.35*	*2.04*	*3.35*	*4.01*	*3.95*	*4.16*	*4.09*
Related Support	0.76	2.26	3.07	3.07	3.72	3.85	4.33
Sales	*9.81*	*9.69*	*10.6*	*11.7*	*11.1*	*11.1*	*10.7*
	6.09	6.48	8.52	10.6	10.2	10.2	9.29
Administrative	*27.0*	*31.1*	*31.1*	*27.2*	*25.1*	*21.9*	*17.9*
Support; Clerical	6.93	6.28	6.43	6.33	7.24	7.47	6.67
Private Household	*1.96*	*3.79*	*0.99*	*0.69*	*1.32*	*1.67*	*1.75*
	0.19	0.36	0.14	0.10	0.15	0.21	0.19
Protective Service	*0.08*	*0.27*	*0.42*	*0.61*	*0.88*	*1.03*	*1.01*
	1.39	2.07	2.32	2.66	2.93	3.25	3.06
Service Except	*18.1*	*13.9*	*14.6*	*13.7*	*13.9*	*16.1*	*17.3*
Protective & Household	4.29	4.46	5.40	6.44	6.48	8.10	8.62
Farming; Fishing;	*3.79*	*1.00*	*0.92*	*0.82*	*0.75*	*0.76*	*1.00*
Forestry	15.9	4.78	3.89	3.38	3.15	3.55	3.40
Precision Production;	*1.93*	*2.18*	*2.28*	*2.09*	*2.24*	*2.12*	*2.32*
Craft; Repair	16.5	21.1	21.3	19.3	18.4	16.3	15.3
Machine Operators;	*18.4*	*13.6*	*8.48*	*5.74*	*4.41*	*2.71*	*2.62*
Assemblers; Inspectors	12.8	11.4	8.89	7.03	6.49	5.22	5.20
Transportation;	*0.49*	*0.52*	*0.52*	*0.59*	*0.65*	*0.64*	*0.58*
Material Moving	6.24	7.13	7.11	6.61	6.72	6.71	8.13
Handlers; Equipment	*0.73*	*1.14*	*1.96*	*1.54*	*1.07*	*0.87*	*1.23*
Cleaners; Laborers	7.31	5.73	5.19	4.94	4.15	4.15	4.89
Military	*0.24*	*0.19*	*0.43*	*0.43*	*0.53*	*0.49*	*0.38*
	4.44	4.49	2.91	2.81	2.73	2.47	2.12

Women's ($\mathrm{NE}_{\mathrm{o},jw,\mathrm{t}}/\mathrm{NE}_{jw,\mathrm{t}}$) italic, Men's ($\mathrm{NE}_{\mathrm{o},jm,\mathrm{t}}/\mathrm{NE}_{jm,\mathrm{t}}$) non-italic, in 10^{-2} (%). NE_o is the national count of people employed in each occupational category, o, as listed. j restricts to age 18-64 non-institutionalized heterosexuals, w to women, m to men, t by year. See equation (38).

Table 1.7: Women's Revealed Preferences

	$\sigma = 0$			$\sigma = 0.5$		
	j = *js*	j = *j*	j = *jl*	j = *js*	j = *j*	j = *jl*
β_w^w	1.770***	1.349***	2.432***	1.768***	1.191***	2.073***
	(0.221)	(0.425)	(0.524)	(0.178)	(0.414)	(0.440)
$\Delta\beta_w^w$	*1.742****	*-0.184*	*-0.205*	*1.203****	*-0.017*	*-0.141*
	(0.318)	*(0.538)*	*(0.668)*	*(0.262)*	*(0.522)*	*(0.585)*
β_w^m	6.630***	7.948***	9.300***	9.397***	8.553***	12.250***
	(0.314)	(0.443)	(0.517)	(0.541)	(0.490)	(0.684)
$\Delta\beta_w^m$	*-1.846****	*-2.399****	*-2.572****	*-3.029****	*-2.927****	*-4.804****
	(0.441)	*(0.557)*	*(0.665)*	*(0.682)*	*(0.597)*	*(0.822)*
AR(1)	0.511	0.600	0.596	0.509	0.602	0.596
Observations	3,798	3,798	3,798	3,798	3,798	3,798

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All with 6 time periods (decades) by 722 location options (Commuting Zones), observation weights proportional to population, regional fixed effects and year by region effects. j restricts to age 18-64 non-institutionalized heterosexuals, s to singles, l to labor force. β_w^w is women's attraction to women's earnings. β_w^m is women's attraction to men's earnings. Italic (Δ) rows give the change over time of each between 1970-1990 and 2000-2019.

Table 1.8: Men's Revealed Preferences

	$\sigma = 0$			$\sigma = 0.5$		
	j = *js*	j = *j*	j = *jl*	j = *js*	j = *j*	j = *jl*
β_m^m	5.903***	7.683***	8.758***	8.247***	8.266***	11.571***
	(0.299)	(0.438)	(0.503)	(0.512)	(0.483)	(0.908)
$\Delta\beta_m^m$	*-1.311****	*-2.233****	*-2.307****	*-2.116****	*-2.733****	*-4.423****
	(0.420)	*(0.549)*	*(0.645)*	*(0.646)*	*(0.588)*	*(0.797)*
β_m^w	1.601***	1.283***	2.213***	1.578***	1.126***	1.846***
	(0.210)	(0.420)	(0.509)	(0.169)	(0.408)	(0.427)
$\Delta\beta_m^w$	*1.648****	*-0.151*	*-0.132*	*1.128****	*0.004*	*-0.052*
	(0.304)	*(0.531)*	*(0.649)*	*(0.249)*	*(0.515)*	*(0.569)*
AR(1)	0.513	0.606	0.612	0.516	0.608	0.611
Observations	3,798	3,798	3,798	3,798	3,798	3,798

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All with 6 time periods (decades) by 722 location options (Commuting Zones), observation weights proportional to population, regional fixed effects and year by region effects. j restricts to age 18-64 non-institutionalized heterosexuals, s to singles, l to labor force. β_m^m is men's attraction to men's earnings. β_m^w is men's attraction to women's earnings. Italic (Δ) rows give the change over time of each between 1970-1990 and 2000-2019.

Table 1.9: Parameter Estimates

	β_w^w	β_w^m	β_m^w	β_m^m	γ_{w-m}^{w-m}	γ_{w+m}^{w-m}	Λ
1970-1990							
$\sigma = 0$	1.35	7.95	1.28	7.68	-0.20	0.33	-0.72
$\sigma = 0.5$	1.19	8.55	1.13	8.23	-0.04	0.06	-0.76
2000-2019							
$\sigma = 0$	1.35	5.55	1.28	5.45	-0.03	0.17	-0.61
$\sigma = 0.5$	1.19	5.63	1.13	5.53	-0.01	0.04	-0.66
Change							
$\sigma = 0$	$0^{\dagger}$	-2.34	$0^{\dagger}$	-2.23	0.17	-0.16	13.8%
$\sigma = 0.5$	$0^{\dagger}$	-2.93	$0^{\dagger}$	-2.73	0.04	-0.02	13.4%

† not statistically significant. β values are from Tables 1.7 and 1.8, for j = j. γ_{w-m}^{w-m} is the elasticity of the women/men (population) sex ratio with respect to the women/men earnings ratio. γ_{w-m}^{w+m} is the elasticity of the women/men sex ratio with respect to earnings overall. Λ quantifies asymmetry of preferences. $\Lambda < 0$ indicates men's earnings are prioritized over women's.

Table 1.10: Effects of Counterfactual UBI on Sex Ratio Imbalance

UBI size $\longrightarrow$		5%	10%	15%	
Median CZ					
$\sigma = 0$	(+1.01%)	1.902%	1.883%	1.854%	(-1.54%)
$\sigma = 0.5$	(+0.80%)	1.898%	1.883%	1.869%	(-0.74%)
Mean CZ					
$\sigma = 0$	(+1.07%)	2.268%	2.244%	2.222%	(-0.98%)
$\sigma = 0.5$	(+0.76%)	2.261%	2.244%	2.228%	(-0.71%)

Sex ratio imbalance $= |\mathrm{N}_{w,\mathrm{z}} - \mathrm{N}_{m,\mathrm{z}}|/(\mathrm{N}_{w,\mathrm{z}} + \mathrm{N}_{m,\mathrm{z}})$, where $\mathrm{N}_{w,\mathrm{z}}$ and $\mathrm{N}_{m,\mathrm{z}}$ are population counts of women and men in Commuting Zone z. Results are for non-institutionalized heterosexuals of age 18-64 (j), in 2000-2019. The middle column is the factual sex ratio imbalance: I assumed 10% de facto post-tax UBI for the model estimation.

1.9 Appendix

1.9.1 Data Sources

As discussed in Section 1.4, I use county level summaries of the Decennial Census and American Community Survey (ACS), and also IPUMS USA microsamples of the same, in concert. I take the county level summaries from Social Explorer,[57] particularly, 1950, 1970, 1980, 1990, and 2000 US Decennial Census summaries, and ACS 2008-2012 and 2015-2019 5-Year Estimate summaries. From IPUMS USA, I take the 1950 1%, 1970 1% metro fm2, 1980 1% metro, 1990 1% unwt, 2000 1% unwt, 2010 ACS 1%, and 2019 ACS 1% samples. I use Dorn's crosswalk files[58] to integrate all of the above data into a panel of 722 Commuting Zones (CZs) by 7 decadal time periods.

At the CZ level, I compute a separate IPUMS measured and Social Explorer (SE) measured version of each variable, as available. In IPUMS, the source observations are of individual persons. Each variable I compute is a weighted average per CZ-year cell, weighted by perwt (person weight) and afactor, as described in Dorn (2009). In SE, the source observations are of counties, which group directly into CZs. For most SE measured variables, I take a population weighted average of the counties in each CZ-year cell.

For every variable at the CZ level, both in IPUMS and in SE, I also calculate versions of the same at two higher levels of aggregation, which I call Sub Region and Super Region. Sub Regions are the 9 regional divisions used by the Census,[59] but split into Urban and Rural CZs within each, for $9 \times 2 = 18$ in total. I define an Urban CZ as one in which, in the 1990 IPUMS sample, there were at least 80,000 people

[57] https://www.socialexplorer.com/explore-tables

[58] See Dorn (2009) and Autor and Dorn (2013).

[59] The divisions are New England (CT, MA, ME, NH, RI, VT), Mid-Atlantic (NJ, NY, PA), East North Central (IL, IN, MI, OH, WI), West North Central (IA, KS, MN, MO, NE, ND, SD), South Atlantic (DC, DE, FL, GA, MD, NC, SC, VA, WV), East South Central (AL, KY, MS, TN), West South Central (AR, LA, OK, TX), Mountain (AZ, CO, ID, MT, NV, NM, UT, WY), and Pacific (CA, OR, WA).

identified as living in an urban area as defined by the Census. Super Regions are the 4 main regions used by the Census, also split into Urban and Rural CZs within each, for 8 in total. The main regions are Northeast (New England, Mid-Atlantic), Midwest (East North Central, West North Central), South (South Atlantic, East South Central, West South Central), and West (Mountain, Pacific).

1.9.2 Data Transformations

I construct final versions of each variable at the CZ level based on weighted averages of several versions of the variable. The weights given to each version are to some extent increasing in the number of individual level observations underlying each. They include IPUMS sourced CZ level, Sub Region level, and Super Region level versions, and transformations of a closely related SE sourced variable at the CZ level, transformed at the Sub Region and Super Region levels. Transformations are of the form,

$$\mathrm{v_Z} = \frac{\mathrm{v}_{\mathrm{Z}}^{IPUMS}}{\mathrm{V}_{\mathrm{Z}}^{IPUMS}} \mathrm{V}_{\mathrm{z}}^{SE}$$

where V is a variable closely related to v that is available through SE, z is CZ, and Z is Region or Super Region. For example, with v as the average earnings of single women, V may be the average earnings of all women.

The use of these weighted averages serves three purposes. First, it guarantees a unique value to every CZ-year cell, for every variable. Second, it uses all available information in a consistent manner, distributing weight in accordance with the amount of information underlying each source estimate, which should improve overall accuracy. Third, the inclusion of Sub Region and Super Region level estimates within each final CZ level estimate accounts for some amount of spillover amongst neighboring CZs, albeit of an arbitrary magnitude. For the average CZ, the IPUMS CZ level estimate accounts for about 80% of the final estimate.

My transformation technique, respective to any given case of v and V, hinges on the assumption that the ratio of v to V in z is equal to that in Z. v and V must be closely related for this assumption to be plausible. A concrete example may be: if women in rural New Hampshire earn 4% more than the average for women in rural New England, then single women in rural New Hampshire are assumed to earn more 4% more than the average for single women in rural New England. Symbolically,

$$\mathrm{v_z} = \delta \mathrm{v_Z} \ , \ \mathrm{V_z} = \delta \mathrm{V_Z}$$

where δ is a scalar, such as 1.04. Therefore,

$$\frac{\mathrm{v_z}}{\mathrm{V_z}} = \frac{\mathrm{v_Z}}{\mathrm{V_Z}}$$

I then assume that SE accurately measures $\mathrm{V_z}$, and IPUMS accurately measures $\mathrm{v_Z}/\mathrm{V_Z}$.

1.9.3 Sexual Orientation

I use a similar data transformation to adjust all measures to capture heterosexuals only. The model is plausible only in application to heterosexuals. Therefore in estimating the model, it is appropriate to refer only to the expected earnings of heterosexuals per se, and migration choices of heterosexuals per se. However, information about sexual orientation has not been traditionally collected in large datasets.

A new question structure in the ACS - codified in the variable SSMC in IPUMS - indicates whether the respondent is in a same sex married couple. For any variable v that is an average over a population count N, let v_M be the same variable restricted to those in married couples, v_{ML} to same sex married couples, v_{MR} to heterosexual couples, and v_R to heterosexuals. Due to SSMC, v_{ML} and N_{ML} are observed, so v_{MR} can be derived algebraically.

$$\mathrm{v}_{MR} = \frac{\mathrm{v}_M \mathrm{N}_M - \mathrm{v}_{ML} \mathrm{N}_{ML}}{\mathrm{N} - \mathrm{N}_{ML}}$$

I then assume that the difference between v_R and v_{MR} is proportional to that between v and v_M,[60]

$$v_R = \frac{v_{MR}}{v_M} v$$

This assumes, for example, that if the average earnings of married men are 8% higher than those of men in general, then the average earnings of married heterosexual men are 8% higher than those of heterosexual men in general. I apply the 2019 values of v_{MR}/v_M across all time periods: but v_{MR}/v_M is specific by all geographies and population subgroups, and v is specific by time period as well.

1.9.4 Calibrations

Two parameters in the model, ζ and σ, are calibrated. Both of these are introduced in Section 1.2.2, though ζ not by name. ζ is responsible for adjusting expected earnings by local cost of living. $\mathring{Y}_z$ denotes expected earnings that have been adjusted both for federal tax as in Appendix section 1.9.6, and for cost of living. Specifically,

$$\mathring{Y}_z = Y_z / R_z^{\zeta}$$
$$y_z = log(\mathring{Y}_z) = log(Y_z) - \zeta \cdot log(R_z)$$

Where R_z is average housing rent in location z, and Y_z has already been adjusted for tax. I calibrate $\zeta = 0.25$. This is the result of regressing $log(Y_z)$ on $log(R_z)$. As such, it captures the extent to which earnings rise with cost of living mechanically.

The calibrated value $\sigma = 0.5$ comes from regressing marriage rates on the sex ratio. This is a relatively high estimate, reflecting the correlation of women's (rather than men's) marriage rates with the sex ratio. Women's marriage rates are more sensitive to the sex ratio than men's are, presumably because women tend to manifest bargaining power by securing marriage matches - see Angrist (2002). Men's bargaining power may manifest in other ways.

[60] This also applies to counts N, with $N_{MR} = N_M - N_{ML}$ in place of v_{MR}, N_M in place of v_M, N in place of v.

1.9.5 Dynamism

As introduced in Section 1.2.1, $\mathrm{u_{j,t}}(\mathrm{x_{t,z}})$ gives the lifetime utility that people expect to arise from their choice to live in any given location in the present time period.[61]

$$\mathrm{u_{j,t}}(\mathrm{x_{t,z}}) = \sum_{t=0}^{\infty} \varphi_{\mathrm{j,t}}^{t} \mathbb{E}_{\mathrm{t}}[\tilde{\mathrm{u}}_{\mathrm{j,t}+t}(\mathrm{x}_{\mathrm{t}+t,z})]$$

where $\varphi_{\mathrm{j,t}} \in [0,1)$ is the discount rate,[62] $\tilde{\mathrm{u}}_{\mathrm{j,t}+t}$ is the flow utility in period $(\mathrm{t} + t)$, and z is a best guess of the optimal choice of location in the future, conditional on the choice to live in z in the present. I make three simplifying assumptions with respect to future time periods. First, I assume that $z = \mathrm{z}$.[63] Second, I assume that people internalize (into their current flow utility) any expectations they have about how their flow utility will change in the future. That is,

$$\mathbb{E}_{\mathrm{t}}[\tilde{\mathrm{u}}_{\mathrm{j,t}+t}(\mathrm{x}_{\mathrm{t}+t,\mathrm{z}})] = \tilde{\mathrm{u}}_{\mathrm{j,t}}(\mathbb{E}_{\mathrm{t}}[\mathrm{x}_{\mathrm{t}+t,\mathrm{z}}])$$

Third, I assume people take location z's present period characteristics as a best guess of its future characteristics. That is,

$$\mathbb{E}_{\mathrm{t}}[\mathrm{x}_{\mathrm{t}+t,\mathrm{z}}] = \mathrm{x_{t,z}}$$

[61] $\mathrm{u_{j,t}}$ evaluates the choice of where to live in the current time period. The choice of where to live in future time periods occurs in future time periods.

[62] $\varphi_{\mathrm{j,t}}$ comprises various forms of uncertainty.

[63] This means that, if the choice to live in z in the present has any effect on what the optimal choice will be in the future, then that effect is only to make z itself more likely to be the optimal choice.

It follows that the lifetime utility that arises from the choice of where to live is scalar multiple of the flow utility, that is,[64]

$$u_{j,t}(x_{t,z}) = \frac{1}{1-\varphi_{j,t}}\tilde{u}_{j,t}(x_{t,z})$$

Therefore, as utility is immune to monotonic transformation, $u_{j,t}$ is the same as the flow utility.

See Bayer et al. (2016) for a more complex treatment of future time periods in the context of migration choice. The main purpose of this complexity, as in Bayer et al., would be correct for some location characteristics $x_{t,z}$ being relatively better predictors of their own future values. (For example, the weather in a Commuting Zone would be a very good predictor of the future weather, while the crime rate would be not as good a predictor of the future crime rate.) However, women's and men's expected earnings are not different in this respect. To test this in my data, I calculate the correlation of Commuting Zone average earnings with lagged earnings (average for the same CZ in the previous time period). For women's earnings, this correlation is 0.89; for men's it is 0.88.

1.9.6 Tax

As referenced in Section 1.2.3, I adjust expected earnings by federal tax and local cost of living. The Census and ACS, via IPUMS USA, provide pre-tax earnings, Y_i, per individual person in the sample, i. National average pre-tax earnings are $\bar{Y}_t = N_t^{-1}\sum_i Y_i$, where N is population count, and t is time period. To arrive at post-tax earnings, I apply the tax function of Heathcote, Storesletten and Violante (2017),

$$\text{post-tax } Y_i \propto (\text{pre-tax } Y_i)^{1-\tau}\ ,\ \tau = 0.181 \tag{22}$$

[64] I assumed geometric discounting and infinite time periods for expositive familiarity, but neither of these are necessary to yield the result that the lifetime utility is a scalar multiple of the flow utility: the discounting could be any function of *t*.

Heathcote et al. show that this function is a close match to the actual federal + state tax and transfer schedule in the US. That is, although the function is fit ($\tau = 0.181$) to capture state taxes (on average) in addition to federal tax, the function itself does not vary by state, so I refer to it throughout the paper simply as federal tax.

In conjunction with (22), it is necessary also to assume the aggregate amount of tax that is collected. I assume total tax is 20% of total pre-tax earnings, that is,

$$N_t^{-1} \sum_i (Y_i - \tilde{\tau} Y_i^{1-\tau}) = 20\% \ \bar{Y}_t \tag{23}$$

$Y_i - \tilde{\tau} Y_i^{1-\tau}$ being the net tax that is collected per individual i, and Y_i being pre-tax earnings. (23) enables algebraic derivation of the constant of proportionality $\tilde{\tau}$, which resolves the exact value of each individual's tax payment $(Y_i - \tilde{\tau} Y_i^{1-\tau})$.

$$\tilde{\tau} = \frac{(1 - 0.2) \sum_i Y_i}{\sum_i Y_i^{1-0.181}} \tag{24}$$

Because $\tau = 0.181 > 0$, this is a progressive tax, taking more in percentage terms from individuals with higher pre-tax earnings Y_i.

Despite being a single function that is applied to all individuals i equally regardless of their location z, the tax (22) does affect the post-tax earnings averages Y_z across locations z differently. Table 1.1, which ranks CZs by their post-tax earnings averages Y_z, would not only list different values if not for the tax, but would be in a different order. For example, San Francisco may have a higher pre-tax average than Southern NJ, but with a relatively more unequal distribution. Because the tax is progressive, it takes a larger share from very high earning individuals, so may take substantially more in tax from San Francisco than it takes from Southern NJ. This could result in San Francisco's post-tax average falling slightly below Southern NJ's, despite the order being the other way around in pre-tax averages.

In addition to paying the tax (22), I assume that people benefit from the tax revenue. Particularly, I assume that half of the tax budget is redistributed equally as in-kind benefits $\underline{\mathrm{Y}}_{\mathrm{t}}$. I add $\underline{\mathrm{Y}}_{\mathrm{t}}$ to each individual's post-tax earnings before calculating the averages Y_{z}. Because the tax budget is 20% of total pre-tax earnings, $\underline{\mathrm{Y}}_{\mathrm{t}}$ is 10%, that is, $\underline{\mathrm{Y}}_{\mathrm{t}} = 10\% \, \bar{\mathrm{Y}}_{\mathrm{t}}$. This, in addition to the progressivity ($\tau > 0$) of the tax payment schedule, further moderates (reduces inequality in) the earnings distribution, both within locations z and across. That is, despite being paid out exactly equally across individuals, $\underline{\mathrm{Y}}_{\mathrm{t}}$ makes a larger difference in relative terms for people with lower earnings.

2 The Distinct Roles of Poverty and Higher Earnings in Motivating Crime

2.1 Introduction

Do the poor have more reason to engage in crime than the rich? And does the presence of higher earnings inspire more crime, by presenting potential offenders with more fruitful opportunities? We model and estimate these distinct forces in a novelly succinct and comprehensive way. As such, this paper articulates the main roles of legitimate (non-crime) earnings in influencing potential criminal offenders' cost-benefit analysis.

Because Poverty[65] operates by making the poor relatively less averse to criminal punishment, it motivates higher rates of crime across the board - both those that are motivated primarily by financial gain to the offender (Financial-Gain crimes),[66] and those motivated primarily by emotional gain (Emotional-Gain crimes).[67] High Earnings, on the other hand, operate by presenting potential offenders[68] with more fruitful opportunities for crime - but fruitful specifically in the sense of financial gain, rather than of emotional gain. These ideas both motivate our simple structural model, and are corroborated in parameter estimates.

The model has two agents - the criminal offender, and the government. The offender's utility weighs his expected gain from crime against his expected loss from associated punishment. The offender's loss coincides with legitimate income that he forgoes when subject to punishment. The government's loss function weighs social

[65]Poverty (proper noun) is a technical term in our model. Specifically, we use the term Poverty to mean average *inverse* earnings, which is a mathematical object that arises in our structural model. In juxtaposition to Poverty, by high or higher Earnings we mean plain average earnings. Although a plain average, average earnings represent the upper tail of the earnings distribution in that earnings are organically right-skewed.

[66]Becker (1968) referred to "monetary" gain versus "psychic" gain. We make this distinction more explicit, and also alter the word choice slightly for modern usage.

[67]Financial-Gain crimes are Robbery, Burglary, Larceny, and Motor Vehicle Theft. Emotional-Gain crimes are Murder, Rape, and Aggravated Assault.

[68]Offenders may be of any income level.

losses from punishment against social losses from crime: due to offenders' response, a higher level of expected punishment per crime leads to a lower level of crime. In equilibrium, the criminal offender maximizes his utility and the government minimizes its loss. The equilibrium crime rate in each Commuting Zone[69] is a function of its average inverse earnings (which we call Poverty), average earnings (Earnings), and controls that include population density and racial demographics.

Criminal punishment plays an essential conceptual role in our model. Although punishment inevitably must play a part in deterring crime, it may be especially relevant in the US context. As shown in Table 2.1, the US has a much higher incarceration rate than any other industrialized country. Many jurisdictions have mandatory sentencing, three strikes laws, capital punishment, and other legislation that leads to high levels of punishment per crime. All of the above implicitly factor in to the government's loss.

Intuitively, earnings influence crime; but the reverse is also true. Where Poverty or Earnings increase, crime increases in response. However, crime may motivate out-migration of people with higher Earnings, or societal trauma that results in more Poverty. Such cases of reverse causality, if unchecked, may undermine the accuracy and precision of estimates. To get around this, we use Shift-Share (Bartik) instruments, for each of Poverty and Earnings. Shift-Share instruments are constructed by interacting non-local (national) averages for each national occupational category with corresponding historical proportions (shares) of each category at the local (Commuting Zone) level. The resulting instruments mitigate reverse causality bias, because local crime rates cannot affect national averages (of earnings or inverse earnings); nor are local crime rates likely to affect historical occupational shares - even at the local (Commuting Zone) level.

We find that all of our main parameter estimates would be biased downward in OLS. Figure 1 provides a simple causal diagram that helps to explain directions of

[69] Commuting Zones are designed to capture local labor markets - see Tolbert and Sizer (1996) and Dorn (2009). We assume they also capture local crime markets.

bias. Poverty motivates crime, and crime can exacerbate Poverty. In other words, the reverse causality has the same sign as the forward causality. Estimated effects of Poverty on crime should be biased downward in OLS in this case. For Earnings, it is the opposite. High Earnings motivate crime, and crime can *decrease* Earnings: OLS may be biased in either direction in this case. Empirically, we find this bias turns out to be downward as well.

One limitation of most literature on the impacts of earnings on crime is that it does not distinguish between the upper and lower tails of the earnings distribution, which may influence crime in different ways. We contribute by modeling both tails of the distribution (Poverty and high Earnings), and their distinct impacts. A second limitation of much past research is the level of jurisdiction covered. The majority of existing literature fails to account for spillover of criminal activity from neighboring areas, or else retreats to the level of states, which are typically too large to represent local markets. We use Commuting Zones as the units of analysis, thereby retaining the level of granularity that is desirable, while also capturing relevant local spillover.

A third limitation is that this literature rarely has clear causal identification. Measured effects on crime are likely to be biased due to reverse causality.[70] We ameliorate this problem by using Shift-Share instruments for each measure of the earnings distribution at the Commuting Zone level. To construct the instruments, we compute earnings distribution summaries for 14 occupational categories at the national level, and take weighted averages of these at the Commuting Zone level - weighted by the categories' shares of the local economy in a lagged time period. The identifying assumption is that local crime rates do not impact the lagged occupational makeup of the economy in each Commuting Zone - nor the national earnings distribution within each occupation - but only the local earnings distribution within each. Thus we isolate the effects of the earnings distribution on crime, sans the effects of crime on earnings.

[70] Gould, Weinberg and Mustard (2002) and Enamorado et al. (2016) use Bartik-like instruments, similarly as we do.

Our parameter estimates indicate that the elasticities of Financial-Gain and Emotional-Gain crime are about the same with respect to Poverty: 0.75 and 0.79, respectively. The more in Poverty a person is, the less legitimate earnings he stands to forego via criminal punishment. This effect applies similarly across both types of crime. The elasticities with respect to high Earnings, on the other hand, are far more disperse: 0.93 and 0.39, respectively. The more Earnings there are present in a Commuting Zone, the larger is the expected benefit to offenders per Financial-Gain crime. The impact of Earnings on Emotional-Gain crime is smaller, and not statistically significant in some specifications. That is, expected benefits per Emotional-Gain crime are not as closely associated with the presence of higher Earnings, because such crimes are not financially motivated.

The remainder of this paper is organized as follows. Section 2.2 covers relevant literature. Section 2.3 describes our data from the Census, American Community Survey, and FBI Uniform Crime Reports. Section 2.4 explains how we scale crime rates, which is vital for making comparisons across different types of crimes. Section 2.5 develops our model. Section 2.6 explains our causal identification in detail. Section 2.7 discusses descriptive statistics to help familiarize the reader with the data. Section 2.8 presents our results. Section 2.9 simulates counterfactual changes in the government's overall level of toughness against crime, exploring the tradeoff that it faces between crime and punishment. Section 2.10 discusses implications, and section 2.11 concludes.

2.2 Related Literature

This paper contributes to the crime literature by highlighting the distinction between Financial-Gain and Emotional-Gain crimes, and most importantly, how lower and higher earnings interact with these two types of crime differently. However, we build on a vast literature of both theoretical and empirical explorations of the economics of crime.

The seminal work of Becker (1968) considers crime in a framework of costs and benefits for the offender, the victim, and the social planner. The offender takes into account the probability of conviction, the punishment if convicted, and other factors when they determine how many offenses to commit within a given time frame. Ehrlich (1970, 1973, 1975) expand upon Becker (1968) and explore the roles that earnings might play in the offender's decision to commit a crime. Witte (1980) builds upon Ehrlich's work and introduces income from illegal activities as a critical factor. Burdett, Lagos and Wright (2003) and Engelhardt, Rocheteau and Rupert (2008) emphasize that crime and income can influence each other and model them as a joint process.

There is a large empirical literature as well. Kelly (2000); Fajnzylber, Lederman and Loayza (2002); Stolzenberg, Eitle and D'alessio (2006); Brush (2007); Choe (2008); Kang (2016); Enamorado et al. (2016) focus on the relationship between inequality and crime. Measures of inequality include the Gini coefficient, the ratio of mean to median household income, and ratio of 90th percentile to 10th percentile. Levels of jurisdiction include country, state, county, and city. Hipp (2007) is noteworthy in examining the role of racial heterogeneity, and also in attempting to distinguish poverty from inequality per se. Closely related to the Financial-Gain mechanism in our paper, Demombynes and Özler (2005) show that wealthier areas see higher Burglary rates.

Most of the papers we have surveyed support a positive relationship between inequality and crime (Fajnzylber, Lederman and Loayza, 2002; Hipp, 2007; Choe, 2008; Kang, 2016; Enamorado et al., 2016). Notably, (Enamorado et al., 2016) estimate a 36% increase in drug-related homicide rate for each one-point increment in the Gini coefficient. However, there is considerable variation in findings. Some authors have found a positive relationship between inequality crime only for certain types of crime, or only under certain conditions (Kelly, 2000; Brush, 2007; Demombynes and Özler, 2005), or occasionally no relationship at all (Stolzenberg, Eitle and D'alessio, 2006).

Our approach may help yield additional insight into this important, yet unsettled topic.

2.3 Data

This study requires a large dataset. Commuting Zone (CZ) level Poverty and Earnings may be either positively or negatively correlated.[71] It follows that large cross sections may be needed to estimate their independent effects. Moreover, each observation in the cross section is a geographical cell - and therefore a large number of individual person observations are needed, separately within every cell, in order to accurately calculate average earnings and other measures within each. This is a challenge because, in most datasets, geographical location is redacted up to a high level of aggregation (such as state, rather than county). And few datasets would have enough observations to occupy a large number of geographical cells, with a large number of observations within each.[72] Finally, in order to constitute a panel, the data must do all of the above for multiple time periods, and in reference to geographical cells that are fixed over time.

IPUMS USA,[73] sourced from the Decennial Census and American Community Survey (ACS), is the largest publicly available compilation of microsamples (samples in which the observations are of individual persons) for the US. IPUMS provides 1% samples of the entire US population. That is, the total number of observations in the sample per time period is about 1% of the national population, so about 3 million (or 2 million, for earlier periods). Observations of individual persons are necessary for flexibility in variable creation. However, we also use Census and ACS summaries at the county level, in addition to IPUMS. In general, IPUMS provides more specific information (variables), while the county level summaries provide more

[71] A positive correlation would be if CZs in which the rich are very rich also tend to be CZs in which the poor are very poor. Alternatively, for a negative correlation, CZs in which the rich are very rich may tend to be CZs in which the poor are only moderately poor in national terms.

[72] See Molloy, Smith and Wozniak (2011).

[73] Ruggles et al. (2020).

complete coverage. We combine them to yield a panel dataset with a combination of specificity and coverage that neither has on its own.

To constitute locations as referenced in the model, it is necessary to use geographical units that meaningfully capture local labor and crime markets. Although Metropolitan Statistical Areas (MSAs) are the most traditional geographical unit for this purpose, Commuting Zones (CZs) have three major advantages.[74] First, CZs are more comparable over time from 1950 to the present. This is because CZs are defined as groups of counties, and county boundaries have remained mostly fixed, while MSA boundaries have been repeatedly adjusted. Second, CZs cover the entire United States, no matter how remote, while MSAs cover only urban areas. Third, CZs are delineated based on actual commuting patterns (albeit in 1990),[75] which means they are ideal for representing actual local markets. We use the set of 722 CZs that are coterminous with the entire continental US (48 states + DC). Each CZ is 4-5 counties on average, so the 722 CZs cover about 3,200 counties in total.

We obtain county level crime data from the FBI's Uniform Crime Report (UCR) detailed arrest and offense reports (Federal Bureau of Investigation, 1980, 1990, 2000, 2010, 2016). Because Commuting Zones (CZs) are defined as particular clusters of counties, county level data aggregates easily into CZ level data. Data from multiple decades enable us to employ a rich set of regional fixed effects to absorb unobserved information that might be correlated with our variables of interest.

We group index crimes into Financial-Gain and Emotional-Gain crimes, which is different than the literature's convention of grouping them into property crime and violent crime. The FBI's definition of property crime includes Burglary, Larceny-theft, Motor Vehicle Theft, and Arson (Federal Bureau of Investigation, 2018*a*). However, though not defined as a property crime, Robbery is also financially motivated, while Arson may not be. Therefore, the FBI's categorization of property crime is not optimal for examining how Poverty and Earnings affect crime. We modify the FBI's cate-

[74] See Dorn (2009) for more discussion.

[75] See Tolbert and Sizer (1996).

gorization by grouping together Robbery, Burglary, Larceny-theft, and Motor Vehicle Theft, and calling the group Financial-Gain crime. Similarly, the FBI's definition of violent crime includes Murder and non-negligent manslaughter, Rape, Robbery, and Aggravated Assault (Federal Bureau of Investigation, 2018*b*). However, Robbery is financially motivated, so it is excluded from our category of Emotional-Gain crime. Therefore, the category of Emotional-Gain crime includes Murder and non-negligent manslaughter, Rape, and Aggravated Assault.

Aggregating crimes at the CZ level has multiple advantages. When residents of one county make regular commutes across county lines, the crimes that they commit or experience are likely to cross county lines as well. If we examine the relationship between Poverty, Earnings and crime at the county level, we may encounter spurious effects of spillovers from one county to another. County borders can be complex, which makes residents of some counties live closer to the epicenters of other counties than to their own. This further increases the possibility of the crime spillover effect. Figure 2 provides an illustrative example. The city of Boston is in Suffolk County, but neighboring Cambridge and Somerville are in Middlesex County. It is not uncommon for residents of Boston to travel to Cambridge and Somerville on a regular basis, and vice versa. Moreover, Brookline is adjacent to Middlesex and Suffolk Counties, but is actually in Norfolk County. Criminals and victims both are likely to be in different counties within their Commuting Zone when crimes occur. Because CZs are defined based on actual commuting flow density, they are more likely to capture a complete picture of crime in the area.

Literature has established that not all serious crimes are reported to the UCR (Skogan, 1975, 1977; Ehrlich, 1977; Myers, 1980; Kennedy, 1988; Coleman and Moynihan, 1996; Levitt, 1998; Baumer and Lauritsen, 2010; Chalfin, 2015). For Financial-Gain crimes, some victims are not interested in reporting the crimes to the police as soon as they recover their items. For Emotional-Gain crimes, some victims may not be inclined to report the crimes, or may not be able to. Our results indicate that Poverty drives higher rates across both types of crime. If crime is under-reported in poorer Commuting Zones, then the actual effects of Poverty would be higher than they appear. This would mean that our estimates constitute a lower bound.

2.4 Standardized Crime Rates

In order to compare rates of crime of different kinds - such as Robbery and Murder - it is necessary to scale these rates relative to one another in a meaningful way. Intuitively, a single Murder is a larger incident than a single Robbery. But how much larger? It may be appropriate to cast the size of each criminal incident in terms of its impact - either in utility to the offender, or in disutility to the victim and society. For example, we might (very roughly) imagine that each Robbery yields a total utility cost to society of $300,000 on average, while each Murder yields an analogous cost of $10 million. However, such notions of impact are subjective, and can only be measured indirectly if at all.

To scale across different types of crime, we rely on the simple fact that more severe crimes tend to occur less frequently than milder crimes. The right-most column of Table 2.2 presents the grand total average rate for each type of crime in our data. The rarest (and intuitively most severe) type of crime is Murder, at 6.8 incidents per 100,000 people per year. Next is Rape, Robbery, Aggravated Assault, Motor Vehicle Theft, Burglary, and finally Larceny - at 2,383 incidents per 100,000 people per year. Following these grand average frequencies, we assume that, for example, a single Murder should be counted similarly as 2,328/6.8 = 350 Larcenies.

Formally, we designate a benchmark type of crime g = G to serve a numeraire for scaling crime rates. We believe Murder is the best choice for this benchmark, in part because typical Murder rates (per 100,000 people per year) fall in the range of 1-10, which is best for scaling purposes. The crime rate $\mathrm{C}_{\mathrm{g,t,z}}$, for any given type of crime g in time period t and location z, is scaled as:

$$\mathrm{C}_{\mathrm{g,t,z}} = (raw\ \mathrm{C}_{\mathrm{g,t,z}}) \cdot \mathrm{C}_G/(raw\ \mathrm{C}_{\mathrm{g}})\ ,\ \mathrm{C}_G = (raw\ \mathrm{C}_G) \tag{25}$$

Each *raw* rate is the actual number of incidents of crime type g per 100,000 people per year. The ratio of grand averages $\mathrm{C}_G/(raw\ \mathrm{C}_{\mathrm{g}})$ serves as a multiplier to make each rate comparable to a Murder rate. The central column of Table 2.2 provides these multipliers. For example, a single Murder is counted as 1 Standardized Incident, while a single Motor Vehicle Theft is counted as only 1/57 of a Standardized Incident.

2.5 Model

Like Becker (1968), we consider the utility of criminal offenders, who weigh their expected gains from crime against their expected losses from associated punishment - and the utility of the government (or Society), which weighs social losses from punishment against social losses from crime. Offenders set the crime rate, and the government sets a punishment policy. Our paper differs in that, where Becker stays in a more abstract space, we use concrete functional forms, which enable us to derive equilibrium solutions and estimate the associated parameters. We also focus much more on the role of expected legitimate (non-crime) earnings as a mechanism of interest, whereas Becker focuses more on risk preferences with respect to expected punishment.

It is worth taking note of the four motives at play in the model. The offender gains from crime and loses from punishment, while the government loses from both crime and from punishment.

	Crime	Punishment
Offender	$\uparrow$	$\downarrow$
Government	$\downarrow$	$\downarrow$

The first three of these motives are highly intuitive. It is the fourth (lower right) that may merit the most contemplation. The government (or Society) suffers from punishment for two reasons. First is simply financial cost - police, prosecution and incarceration cost money. Second, more deeply, is that the government internalizes (in principle) total social welfare, which includes the welfare of the prosecuted, both guilty and innocent. Note that this is entirely sufficient to yield, as a corollary, the intuitive idea that the innocent should not be prosecuted (or convicted). Because prosecution cannot deter crime except in so far as it is in response to crime, prosecution of the innocent has no useful function, and is pure loss. Prosecution of the guilty, though it may even be equally painful as prosecution of the innocent, can be justified only because it has a useful function as deterrent.

We begin by interpreting Becker's offender's utility in a more concrete functional form. An offender's expected net gain from criminal activity can be written as:

$$U(C_i) = \frac{B}{1-\alpha} C_i^{1-\alpha} - KC_i \ , \ \alpha \in [0,1] \tag{26}$$

where i is the offender (individual person), C_i is i's chosen frequency of crime, B is the expected benefit (gain to the offender) per crime, and K is the expected punishment per crime. We assume that the offender is risk-neutral with respect to punishment,[76] and use the power $(1-\alpha) \in [0,1]$ to represent Becker's supposition that offenders face decreasing marginal gain from crime.

We make two innovations upon equation (26). First, we assume that offenders, in addition to facing decreasing marginal gain from crime, face increasing marginal punishment:

$$U(C_i) = \frac{B}{1-\alpha} C_i^{1-\alpha} - \frac{K}{1+\tilde{\alpha}} C_i^{1+\tilde{\alpha}}$$

For convenience, we suppose particularly that $\tilde{\alpha} = (1-\alpha)$:[77]

$$U(C_i) = \frac{B}{1-\alpha} C_i^{1-\alpha} - \frac{K}{2-\alpha} C_i^{2-\alpha} \ , \ \alpha \in [0,1]$$

The rationale for increasing marginal punishment is that a more frequent offender is more likely to catch the focus of law enforcement and prosecution. Second, we assume that the offender's loss from criminal punishment is proportional to his expected legitimate (non-crime) earnings Y_i (plus a baseline $\underline{Y}$):[78]

$$U(C_i) = \frac{B}{1-\alpha} C_i^{1-\alpha} - \frac{K}{2-\alpha} C_i^{2-\alpha} (Y_i + \underline{Y}) \ , \ \alpha \in [0,1] \tag{27}$$

The rationale for this critical assumption is that people with higher expected earnings stand to lose more from criminal conviction.[79] In the case of incarceration, the offender would lose any earnings he would otherwise have gained if not behind bars. Even without incarceration, criminal records would still carry severe reputational damage, likely to impair the offender's future career. Note however that we leave the benefit B (as opposed to the cost K) portion of the utility function independent

[76] Becker (1968) focuses largely on the distinction between the probability of conviction, and the severity of punishment conditional on condition; but this is not important for purposes. Our assumption of risk-neutrality renders the distinction inconsequential: K is the overall expected punishment per crime, a composite of the expected probability and the expected conditional punishment.

[77] In the limiting case of $\alpha = 1$, this is $U(C_i) = B \cdot log(C_i) - K \cdot C_i$.

[78] We assume $\underline{Y}$ is equal to 10% of average pre-tax earnings - see Appendix section 2.13.3.

[79] Wealthier offenders' ability to pay for better lawyers may dampen this effect, but cannot negate or reverse it.

of earnings. This independence can be questioned on the grounds that offenders may have decreasing marginal utility in total (legitimate + criminal) earnings, and therefore that criminal earnings may mean more in utility terms to lower legitimate-earning offenders. However, higher earning offenders may also have access to higher quality (such as “white-collar") opportunities for crime. Implicitly, we assume that these two factors offset one another.

The second agent in the model is the government, which chooses the level of punishment per crime K. (K is a composite of the probability of conviction, and the expected punishment conditional on conviction.) We assume that the government faces a loss function in which it suffers both from higher levels of K, and from higher crime rates C:[80]

$$\mathrm{L}(\mathrm{K}) = \vartheta^{-1}\mathrm{K}^{\vartheta} + \mathrm{C}(\mathrm{K}) \ , \ \vartheta > 0 \tag{28}$$

Higher punishment per crime K is a loss to society for multiple reasons. Punishment is detrimental to offenders, whose welfare is part of the social welfare. Higher punishment per crime also implies costly government activities - more police, more prosecution, more incarceration - all of which fall heavily on the taxpayer. As the model is concerned, it is important only that higher K is undesirable.

Our analysis concerns different types of crime g, over different Commuting Zones z and time periods t. Written fully, equations (27) and (28) are:

Offender's Gain:

$$\mathrm{U}(\mathrm{C}_{\mathrm{i,g,t,z}}) = \frac{\mathrm{B}_{\mathrm{g,t,z}}}{1-\alpha_{\mathrm{g}}}\mathrm{C}_{\mathrm{i,g,t,z}}^{1-\alpha_{\mathrm{g}}} - \frac{\mathrm{K}_{\mathrm{g,t,z}}}{2-\alpha_{\mathrm{g},}}\mathrm{C}_{\mathrm{i,g,t,z}}^{2-\alpha_{\mathrm{g}}}(\mathrm{Y}_{\mathrm{i,t,z}} + \underline{\mathrm{Y}}_{\mathrm{t}}) \ , \ \alpha_{\mathrm{g}} \in [0,1] \tag{29}$$

Government's Loss:

$$\mathrm{L}(\mathrm{K}_{\mathrm{g,t,z}}) = \vartheta_{\mathrm{g}}^{-1}\mathrm{K}_{\mathrm{g,t,z}}^{\vartheta_{\mathrm{g}}} + \mathrm{C}(\mathrm{K}_{\mathrm{g,t,z}}) \ , \ \vartheta_{\mathrm{g}} > 0 \tag{30}$$

[80]Crime rates C are expressed in Standardized Incidents per 100,000 people per year - see Table 2.2.

The optimality conditions for each (given $\mathrm{U}''(\mathrm{C}^*_{\mathrm{i,g,t,z}}) < 0$, $\mathrm{L}''_{\mathrm{t}}(\mathrm{K}^*_{\mathrm{g,t,z}}) > 0$)[81] are:

$$\mathrm{U}'(\mathrm{C}_{\mathrm{i,g,t,z}}) = 0 \ , \ \mathrm{L}'(\mathrm{K}_{\mathrm{g,t,z}}) = 0$$

That is, the offender i chooses a crime frequency $\mathrm{C}_{\mathrm{i,g,t,z}}$ to maximize U, and the government chooses an expected level of punishment per crime $\mathrm{K}_{\mathrm{g,t,z}}$ to minimize L.

For the offender, $\mathrm{U}'(\mathrm{C}_{\mathrm{i,g,t,z}}) = 0$ is equivalent to:

$$\mathrm{C}^*_{\mathrm{i,g,t,z}} = \mathrm{M}_{\mathrm{i,t,z}} \frac{\mathrm{B}_{\mathrm{g,t,z}}}{\mathrm{K}_{\mathrm{g,t,z}}} \ , \ \mathrm{M}_{\mathrm{i,t,z}} = (\mathrm{Y}_{\mathrm{i,t,z}} + \underline{\mathrm{Y}}_{\mathrm{t}})^{-1} \tag{31}$$

The crime rate (in each Commuting Zone z in each time period t) then follows as the average crime frequency over individuals i:

Crime Rate:

$$\mathrm{C}_{\mathrm{g,t,z}} = \mathrm{N}^{-1}_{\mathrm{t,z}} \sum_{\mathrm{i} \in \mathrm{t,z}} \mathrm{C}^*_{\mathrm{i,g,t,z}} = \frac{\mathrm{M}_{\mathrm{t,z}} \mathrm{B}_{\mathrm{g,t,z}}}{\mathrm{K}_{\mathrm{g,t,z}}} \ , \ \mathrm{M}_{\mathrm{t,z}} = \mathrm{N}^{-1}_{\mathrm{t,z}} \sum_{\mathrm{i} \in \mathrm{t,z}} (\mathrm{Y}_{\mathrm{i,t,z}} + \underline{\mathrm{Y}}_{\mathrm{t}})^{-1} \tag{32}$$

Plugging this in to (30), the government's loss function becomes:

$$\mathrm{L}(\mathrm{K}_{\mathrm{g,t,z}}) = \vartheta_{\mathrm{g}}^{-1} \mathrm{K}^{\vartheta_{\mathrm{g}}}_{\mathrm{g,t,z}} + \frac{\mathrm{M}_{\mathrm{t,z}} \mathrm{B}_{\mathrm{g,t,z}}}{\mathrm{K}_{\mathrm{g,t,z}}} \tag{33}$$

The government's optimality condition ($\mathrm{L}'(\mathrm{K}_{\mathrm{g,t,z}}) = 0$) therefore is equivalent to:

$$\mathrm{K}^*_{\mathrm{g,t,z}} = (\mathrm{M}_{\mathrm{t,z}} \mathrm{B}_{\mathrm{g,t,z}})^{1/(1+\vartheta_{\mathrm{g}})} \tag{34}$$

Plugging this back in to (32), the equilibrium crime rate is:

$$\mathrm{C}_{\mathrm{g,t,z}} = \frac{\mathrm{M}_{\mathrm{t,z}} \mathrm{B}_{\mathrm{g,t,z}}}{(\mathrm{M}_{\mathrm{t,z}} \mathrm{B}_{\mathrm{g,t,z}})^{1/(1+\vartheta_{\mathrm{g}})}} = (\mathrm{M}_{\mathrm{t,z}} \mathrm{B}_{\mathrm{g,t,z}})^{\vartheta_{\mathrm{g}}/(1+\vartheta_{\mathrm{g}})} \tag{35}$$

[81] See Appendix section 2.13.4 for derivation.

This result is intuitive. The equilibrium value of punishment per crime $K_{g,t,z}$ increases in each Commuting Zone z in response to the crime rate in each,[82] but without overcompensating. Therefore, punishment acts simply to dampen the crime rate. If punishment were constant over Commuting Zones, then the crime rate would be $C_{g,t,z} = M_{t,z} B_{g,t,z}$. Instead, due to punishment's response to the crime rate, the latter is dampened by the power $\vartheta_g/(1 + \vartheta_g)$, which is necessarily in the interval (0, 1) for $\vartheta_g > 0$.

We assume that the expected gain to offenders per crime $B_{g,t,z}$ is a Cobb-Douglas function of average earnings $Y_{t,z}$, and controls $X_{t,z}$ that include population density and racial demographics:

$$B_{g,t,z} = Y_{t,z}^{\beta_g^y} \cdot \{X_{t,z}^{\beta_g^x}\} \ , \ Y_{t,z} = N_{t,z}^{-1} \sum_{i \in t,z} (Y_{i,t,z} + \underline{Y}_t) \tag{36}$$

The expected gain per crime $B_{t,z}$ is increasing in average earnings for the reason that higher earning people, as potential victims, possess more that is worth stealing. It is increasing in population density for the reason that higher interpersonal contact rates will present offenders with a more plentiful flow of potential crime opportunities, some of which will be better than others.

As additional controls $\{X_{t,z}\}$ we take, in the Baseline specification, the fractions of the local population who are Black and White, as well as regional fixed effects, and year by region effects. These Baseline controls are those which we are most confident to be exogenous: racial fractions are deeply inertial and persistent, because race is fully genetic, and embedded in particular locations by family ties. In a Fully Controlled specification, we include an additional five controls that we are less confident to be exogenous: the fractions of the local population who are Young (under 40), Jailed (currently incarcerated), College educated, Married, and Male. All of these demographic fractions should be interpreted (under the model) as forces that exert

[82] $K_{g,t,z}$ may be raised via heightened policy activity, or more aggressive prosecution.

cultural influences on the perceived value of criminal activity, to all potential offenders in each location.

Plugging (36) into (35) and taking logs, we arrive at a closed-form linear expression for the equilibrium log crime rate with respect to observed variables:

Equilibrium:

$$c_{g,t,z} = \Theta_g m_{t,z} + \Theta_g \beta_g^y y_{t,z} + \{\Theta_g \beta_g^x x_{t,z}\} \tag{37}$$

$$c_{g,t,z} = \mathit{log}(C_{g,t,z}) \ , \ m_{t,z} = \mathit{log}(M_{t,z}) \ , \ y_{t,z} = \mathit{log}(Y_{t,z}) \ , \ x_{t,z} = \mathit{log}(X_{t,z})$$

$$\Theta_g = \vartheta_g/(1+\vartheta_g)$$

where $Y_{t,z}$ is the average earnings (in Commuting Zone z in time period t), $M_{t,z}$ is the average inverse earnings, and $X_{t,z}$ are controls, including population density and racial demographics. We estimate this equation using Two-Stage Least Squares, instrumenting for $m_{t,z}$ and $y_{t,z}$ with Bartik instruments. The coefficient on $m_{t,z}$ recovers $\Theta_g = \vartheta_g/(1+\vartheta_g)$. The coefficient on $y_{t,z}$ recovers $\Theta_g\beta_g^y$, which implies the value of β_g^y given that for Θ_g; and so on for each β_g^x.

$Y_{t,z}$ and $M_{t,z}$ are the explanatory variables of interest. $Y_{t,z}$ represents the upper tail of the earnings distribution, while $M_{t,z}$ represents poverty (the lower tail). Although $M_{t,z}$ is the average inverse of earnings, it is *not* the inverse of $Y_{t,z}$.

$$Y_{t,z} = N_{t,z}^{-1} \sum_{i \in t,z} (Y_{i,t,z} + \underline{Y}_t)$$

$$M_{t,z} = N_{t,z}^{-1} \sum_{i \in t,z} (Y_{i,t,z} + \underline{Y}_t)^{-1} > Y_{t,z}^{-1}$$

The latter inequality holds unless all values of $Y_{i,t,z}$ are identical, in which case $M_{t,z} = Y_{t,z}^{-1}$. Given any variation in individual earnings ($Y_{i,t,z}$), then $M_{t,z} > Y_{t,z}^{-1}$, because $M_{t,z}$ places disproportionate weight on the lower tail of the distribution. As such $M_{t,z}$, which arises organically as a central variable in the model from the assumption that lower earning people have less to lose in the context of criminal punishment, is ideal as a measure of poverty. $Y_{t,z}$ is merely the average, but places disproportionate weight on the upper tail in the sense that it is affected by the right-skew characteristic of earnings distributions.[83] Thus, $Y_{t,z}$ and $M_{t,z}$ capture the upper and lower tails in juxtaposition.

2.6 Identification

Two main identification hurdles arise with regard to estimating (37): reverse causality, and omitted explanatory variables. Our responses to each in turn are: Shift-Share (Bartik) instruments, and controls that include regional fixed effects and year by region effects.

2.6.1 Shift-Share (Bartik) Instruments

In equation (37), the response (left hand side) variables are crime rates, while the main explanatory (right hand side) variables of interest are average earnings, and average inverse earnings. The basic economic idea coinciding with equation (37) is that the distribution of legitimate (non-crime) earnings plays a role (or multiple roles) in motivating people to engage in crime. However, crime rates may also have reverse effects on the distribution of legitimate earnings. Unfavorable crime rates may motivate individuals with the means to move to do so, thereby dampening what would otherwise have the upper tail of the earnings distribution in any given Commuting Zone. Amongst the poor, crime may contribute to negative feedback loops, pushing more people into crime.

To get around any such effects of crime on earnings, we construct a set of Shift-Share (Bartik) instruments.[84] The instrument, for each variable, interacts the lagged

[83] $Y_{t,z}$ places disproportionate weight on the upper tail relative to, for example, the median of $Y_{i,t,z}$, or the average of $log(Y_{i,t,z})$.

[84] The name Bartik comes from Bartik (1991), but refers to a broad class of instruments. See Bound and Holzer (2000) and Diamond (2016) for examples.

fractional breakdown of employment opportunities in each location z across occupational categories, with the corresponding vector of national averages for each category. This arrives at a measure for each earnings variable that is decoupled from the effects of contemporary local crime rates in each location z, depending rather only on the historical structure of the economy in each.

The component of the instrument that is specific to each location z is what we call the Occupational Profile (OP). The OP of each Commuting Zone z, in each time period t, is a vector of 14 fractions summing to one, giving the occupational breakdown of employment across each in z. That is, the OP for t,z is the vector,

$$\left\{\frac{\mathrm{NE}_{\mathrm{o,t,z}}}{\mathrm{NE}_{\mathrm{t,z}}}\right\}_{\mathrm{o}}$$

where $\mathrm{NE}_{\mathrm{o,t,z}}$ is the count of people employed in Occupation o, and $\mathrm{NE}_{\mathrm{t,z}}$ is the count of people employed in any Occupation. To arrive at the instrument for Earnings, the OP is interacted with the vector,

$$\left\{\mathrm{Y}_{\mathrm{o,t}}\right\}_{\mathrm{o}}$$

where $\mathrm{Y}_{\mathrm{o,t}}$ is the national average post-tax earnings of people who identify Occupation o as their habitual occupation (even if not employed).

We assume that OPs are a first mover of the economic system under study, arising from a combination of exogenous geographical, historical, and technological factors. In other words, OPs do not depend on crime rates. For example: although unfavorable crime rates may influence average earnings in Chicago via out-migration or societal trauma, crime does not influence the probability that a worker in Chicago, if employed, will be employed in finance. In case this is considered questionable, we lag OPs by one time period. The one period lag allows OPs to update for relevance to the nature of economic activity in the current time period, while remaining clearly causally prior to crime in the current period. The instrument for Earnings therefore is,

$$\tilde{\mathrm{y}}_{\mathrm{t,z}} = log(\tilde{\mathrm{Y}}_{\mathrm{t,z}}) \ , \ \tilde{\mathrm{Y}}_{\mathrm{t,z}} = \sum_{\mathrm{o}} \frac{\mathrm{NE}_{\mathrm{o,t-1,z}}}{\mathrm{NE}_{\mathrm{t-1,z}}} \cdot \mathrm{Y}_{\mathrm{o,t}} \tag{38}$$

where t = (0, 1, 2, 3, 4) indicate survey years (1980, 1990, 2000, 2010, 2016) respectively. The instrument for Poverty is analogous,

$$\tilde{\mathrm{m}}_{\mathrm{t,z}} = log(\tilde{\mathrm{M}}_{\mathrm{t,z}}) \ , \ \tilde{\mathrm{M}}_{\mathrm{t,z}} = \sum_{\mathrm{o}} \frac{\mathrm{NE}_{\mathrm{o,t-1,z}}}{\mathrm{NE}_{\mathrm{t-1,z}}} \cdot \mathrm{M}_{\mathrm{o,t}} \tag{39}$$

The instrument given by (39) is novel, but follows by the same logic as (38), which is a more standard Shift-Share instrument. $\mathrm{M}_{\mathrm{o,t}}$, like $\mathrm{Y}_{\mathrm{o,t}}$, is a summary of the distribution of Occupation o specific (but not location z specific) earnings at the national level, so is exogenous for the same reasons.

2.6.2 Fixed Effects

As is almost always the case in any real setting, the potential for omitted unobservables to yield bias is the truest and most intractable identification issue. We rely on a standard method that is meant to (plausibly/perhaps) absorb all relevant unobserved information, namely, fixed effects. Because not all potentially relevant factors are of a nature that is fixed over time, we use region-by-year effects as well regional fixed effects. For example, the weather in New York is about equally worse than the weather in Florida regardless of the time period, but the same cannot always be said of things like cultural influences. We use the nine regional divisions used by the Census,[85] and further split each of these into their urban and rural Commuting Zones,[86] for a total of 18 regions. We include an effect for each of these, in each of the 5 decadal time periods, for a total of 90 effects.

[85]These nine regional divisions have the benefit of being about the same size as one another.
[86]See Appendix section 2.13.1.

2.7 Descriptive Statistics

This paper uses a host of socioeconomic and demographic factors as explanatory variables in determining equilibrium crime rates. Table 2.4 summarizes these explanatory and control variables, scaled for readability. Means are in rows without parentheses, while the standard deviation for each is in parentheses in the row beneath. The percentage of Whites in the mean Commuting Zone has decreased over the course of 1980-2016, as has the percentage of Young (under 40) people, and the Marriage rate. Meanwhile, the percentage of Blacks, the incarceration (Jailed) rate, and the percentage of people who are College educated in the mean CZ have all increased. All of the above reflect, but are not equivalent to, corresponding changes in the national mean of each variable.[87]

It should be noted that although mean values of all the variables as discussed in the previous paragraph have changed over time, this does not make a difference in the model estimation. In the same vein, adjusting earnings figures for inflation would not make any difference. This is because, in logarithmic form, all scaling factors become additive intercepts, which in the model estimation are absorbed in time period effects fixed effects.

Unlike the scaling of variables above, the scaling of crime rates is extremely important. Table 2.3 summarizes the distribution of (CZ average) crime rates, expressed in Standardized Incidents as defined in Section 2.4 and Table 2.2. Financial-Gain crime is Robbery, Burglary, Larceny, and Motor Vehicle Theft. Emotional-Gain crime is Murder, Rape, and Aggravated Assault. Both types of crime peaked in 1990. More concretely still, Table 2.5 lists individual Commuting Zones by their crime rates in 2010. One illustrative example is Detroit - 7th in Emotional-Gain crimes, but not in the top 15 for Financial-Gain crimes. Detroit has a pronounced lower tail of earnings (Poverty) - so has high crime rates in general. But it does not have a pronounced

[87]The national mean would weight each individual equally. The values in Table 2.4 are means of CZ means, weighting each CZ equally, which in effect gives more weight to individuals in lower populated CZs.

upper tail of earnings. Therefore in Financial-Gain crimes - which are driven by both Poverty and high Earnings - Detroit is overtaken by some higher earning CZs such as Houston and San Francisco.

2.8 Results

Our main parameters of interest are ϑ_{g} and β^{y}_{g}, for each of Emotional-Gain crimes ($\mathrm{g} = E$) and Financial-Gain crimes ($\mathrm{g} = F$). ϑ identifies the role of Poverty in driving crime,[88] while β^{y} identifies the role of high Earnings. Our central finding, both theoretically and empirically, is that β^{y} is higher for Financial-Gain crimes than it is for Emotional-Gain crimes ($\beta^{y}_{F} > \beta^{y}_{E}$), while ϑ is about the same for both types of crime ($\vartheta_{F} = \vartheta_{E}$). Table 2.8 encapsulates these main results. Additionally, we find that not accounting for reverse causality would underestimate both parameters, for both types of crime, often by more than 50%.[89]

The distinct roles of Poverty and higher Earnings in driving crime are such that the former motivates all types of crime, whereas the latter motivates only (or principally) crimes coinciding with financial gain to the offender. The role of Poverty enters into the model via the assumption that the offender's disutility in the event of criminal punishment is proportional to his legitimate (non-crime) earnings.[90] This assumption, combined with optimality and equilibrium conditions, yields the average inverse of earnings (the variable $\mathrm{m}_{\mathrm{t,z}}$) as a structural representative of Poverty. Therefore the parameter, ϑ, that is associated with $\mathrm{m}_{\mathrm{t,z}}$ should attain similar estimates across both categories of crime (Financial-Gain and Emotional-Gain).

The upper tail of the earnings distribution, unlike Poverty (the lower tail), should not motivate both types of crime commensurately. High Earnings enter the model via the assumption that higher earning people possess more that is worth stealing,

[88] ϑ also pins down the government's loss function, by weighing the harm arising from higher punishment per crime against the harm arising from crime directly.

[89] This is apparent in Table 2.7.

[90] That is, people with lower legitimate earnings have less to lose the context of criminal punishment, and therefore more reason (on balance) to engage in crime.

and therefore present more fruitful criminal opportunities to potential offenders. This logic applies far more to Financial-Gain crimes (Robbery, Burglary, Larceny, Motor Vehicle Theft) than it does to Emotional-Gain crimes (Murder, Rape, Aggravated Assault). Therefore the parameter β^y, which captures the role of high Earnings, should be higher for the Financial-Gain crimes than it is for Emotional-Gain crimes.

The parameter estimates summarized in Table 2.8 corroborate that the role of high Earnings (β^y) is higher for Financial-Gain crimes (1.185) than for Emotional-Gain crimes (0.525), while the role of Poverty, $\Theta = \vartheta/(1+\vartheta)$, is about the same (0.788 versus 0.750) across both categories of crime. These estimates are taken from Table 2.7's lower panel, the Fully Controlled specification. The Baseline specification (Table 2.7's upper panel) suggests that β^y for Emotional-Gain crimes is not statistically significant at all, which would corroborate our theory even more strongly.

One reason that β^y may be positive for Emotional-Gain crimes (albeit not as large as that for Financial-Gain crimes) is that the two types of crime are linked via gang activity. Murder, for example, although not intrinsically associated with financial gain for the offender, may be deployed strategically by criminal enterprises that are driven ultimately by illegal revenue streams. A second reason is that the presence of high Earnings may inspire envy amongst potential offenders, thus enhancing the emotional impetus for crime.

2.9 Counterfactual Toughness on Crime

The most important parameter in our model, ϑ, is also one for which we receive some of the most consistent estimates across different specifications. By definition, ϑ is the government's distaste for criminal punishment: a lower value of ϑ indicates a tougher stance on crime. By derivation, ϑ also coincides with the elasticity of crime with respect to Poverty; and is a factor in the elasticity of crime with respect to all other motivating factors as well.

Because ϑ is a parameter of the government's loss function, it is in essence a policy choice. We simulate counterfactual changes in the value of ϑ in order to examine the tradeoff it faces between crime and punishment. It follows[91] from the equilibrium crime rate (35) and equilibrium level of expected punishment per crime (34) that,

$$\check{\check{\mathrm{C}}}_{\mathrm{Z}} = \mathrm{C}_{\mathrm{Z}}^{\check{\check{\vartheta}}/(1+\vartheta(1+\check{\check{\vartheta}}))} - 1$$

$$\check{\check{\mathrm{K}}}_{\mathrm{Z}} = \mathrm{C}_{\mathrm{Z}}^{-\check{\check{\vartheta}}/(1+\vartheta(1+\check{\check{\vartheta}}))} - 1$$

$$\check{\check{\vartheta}} := (\check{\vartheta} - \vartheta)/\vartheta \ , \ \check{\check{\mathrm{C}}}_{\mathrm{Z}} := (\check{\mathrm{C}}_{\mathrm{Z}} - \mathrm{C}_{\mathrm{Z}})/\mathrm{C}_{\mathrm{Z}} \ , \ \check{\check{\mathrm{K}}}_{\mathrm{Z}} := (\check{\mathrm{K}}_{\mathrm{Z}} - \mathrm{K}_{\mathrm{Z}})/\mathrm{K}_{\mathrm{Z}}$$

where $\check{\vartheta}$ is a counterfactual theta value of policy choice, and $\check{\mathrm{C}}_{\mathrm{Z}}$ and $\check{\mathrm{K}}_{\mathrm{Z}}$ are the resulting counterfactual equilibrium crime rate and level of expected punishment per crime, respectively. For example, a value of $\check{\check{\vartheta}} = 0.5$ would mean that the counterfactual policy choice is to increase ϑ by 50%. $\check{\check{\mathrm{K}}}_{\mathrm{Z}} = -0.3$ would mean that the resulting equilibrium expected punishment per crime would be 30% lower than in the factual equilibrium, and $\check{\check{\mathrm{C}}}_{\mathrm{Z}} = 0.2$ that the resulting crime rate would be 20% higher.

As ϑ represents the government's distaste for punishment, a higher value of ϑ must coincide with a loosening of punishment, and consequently higher equilibrium crime rates. The two right columns of Table 2.9 summarize the consequences of a 20% increase in ϑ_{g} for each type of crime g. For Financial-Gain crimes ($\mathrm{g} = F$), this would result in a 6.2% decrease in expected punishment per crime in the median Commuting Zone, and a corresponding 6.6% increase in crime. A 20% decrease in ϑ, on the other hand, would result in 9.2% increase in expected punishment per crime, and a 8.4% decrease in crime.

Being an abstract policy disposition, it is not obvious what a 20% change in the value of ϑ would look like concretely. However, it links together two concrete outcomes - the crime rate (C), and expected level of punishment per crime (K). A 9.2% increase in K can be interpreted either as a 9.2% increase in the probability of conviction, or a 9.2% increase in the severity of punishment conditional on conviction, with the other

[91] See Appendix section 2.13.5 for derivation.

held constant. Our results indicate that, for example, in order to achieve an 8.4% decrease in crime, the median Commuting Zone would need to increase the severity of its criminal sentences (or its conviction probability) by 9.2%. In other words, there are no particularly great options in the tradeoff between crime and punishment.

2.10 Discussion

Our results provide concrete evidence for the benefits of reducing inequality, and especially poverty. Many current poverty alleviation efforts exist in the US. For example, summer youth employment programs (SYEP) are known to alleviate both poverty and crime (Modestino, 2019; Davis and Heller, 2020; Kessler et al., 2021). Researchers have found that SYEP increase employment outcomes in a subset of youths. They have also found suggestive evidence that the programs improved youths' conflict resolution skills, including self-regulation and ability to respond positively to criticism. Additional improvements may include peer networks and income. Currently these programs are conducted in Boston, Chicago, and New York: more widespread adoption would certainly be beneficial. Another method is Moving to Opportunity (MTO) (Chetty, Hendren and Katz, 2016). Researchers found that MTO helps children before age 13 in areas including college attendance, earnings, and single parenthood rates.

Another important line of work highlights the vicious cycle between poverty and productivity decline (Banerjee and Mullainathan, 2008; Kaur et al., 2021; Duquennois, 2022). Increasing the savings rate of low-income individuals can mitigate this productivity decline. There are two obstacles to this solution. First, low-income individuals tend to be more present-biased before payday (Carvalho, Meier and Wang, 2016). Second, low-income households have low participation rate in 401(k) (Poterba, Venti and Wise, 2000). To overcome these problems, companies should automatically enroll their low-income employees in 401(k) (Thaler and Benartzi, 2004). Bhargava and Manoli (2015) have found mailing, simplification, and heightening salience of benefits increase take-up of EITC benefits. Hoynes, Schanzenbach and Almond (2016) have found that access to food stamps in childhood increases health outcomes and, for women, later-life economic self-sufficiency.

Finally, taxes can play a large role in social welfare. A negative tax on the poor can help overcome hyperbolic discounting and myopia (Farhi and Gabaix, 2020). Other researchers have suggested overall tax reforms to reduce inequality (Ales, Kurnaz and Sleet, 2015; Altig and Carlstrom, 1999). The above methods of reducing inequality and poverty, amongst others, can increase legitimate earnings of the poor and in turn reduce crime.

2.11 Conclusion

Much work in economics has argued that earnings inequality plays a role in driving crime. This paper is novel in distinguishing the role of the lower tail of the earnings distribution (Poverty) in driving crime, from that of the upper tail. We find that the roles played by these two tails driving crime are different in kind. Poverty results in individual offenders who have less to lose from criminal punishment, and are therefore less averse to engaging in all forms of crime. High Earnings heighten the expected benefit to offenders per crime, but only for crimes that yield financial gain to the offender. As a result, high Earnings drive only a subset of crimes, while Poverty results in higher rates of all types of crime.

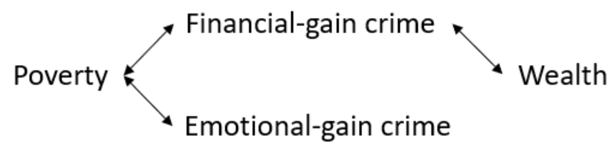
We develop a new model that articulates how Poverty and Earnings become factors in determining equilibrium crime rates. The model has two players - the criminal offender, and the government. The offender maximizes his utility by choosing his frequency of crime. The government minimizes social loss by choosing the level of expected punishment per crime. In equilibrium, the crime rate in each Commuting Zone is a function of Poverty, Earnings, and demographic factors.

In order to estimate the model, we construct a comprehensive panel of Census, American Community Survey, and FBI Uniform Crime Reporting data, covering the entire United States (722 Commuting Zones) from 1980-2016. To correct for reverse causality, we deploy novel Shift-Share instruments for different parts of the earnings distribution. The instruments assume that crime impacts within-industry earnings in each Commuting Zone, but not the historical industrial makeup of each CZ, nor within-industry earnings distributions at the national level.

We find that higher Earnings significantly increase rates of Financial-Gain crimes (Robbery, Burglary, Larceny, Motor Vehicle Theft), while Poverty significantly increases rates both of Financial-Gain and of Emotional-Gain crimes (Murder, Rape, Aggravated Assault). Removing reverse causality results in substantially higher parameter estimates, in some cases by more than double.

2.12 Tables and Figures

Figure 1: Variable Relationsips



Poverty motivates (+) crime, but crime can exacerbate (+) Poverty. OLS results are therefore biased downward. High Earnings motivate (+) Financial-Gain crime, and crime can diminish (–) Earnings, such as via out-migration. The resulting direction of bias in OLS can be either upward or downward for this latter case.

Table 2.1: Incarceration Rates in Industrialized Countries

Country	Incarceration Rate
United States	642
New Zealand	214
Australia	172
United Kingdom	139
Spain	128
Portugal	125
France	114
Canada	112
Italy	99
Austria	98
Greece	93
Belgium	90
Switzerland	82
Ireland	80
Germany	77
Denmark	65
Norway	65
Netherlands	63
Sweden	63
Finland	53
Japan	40

Rates are per 100,000 population in 2017-2018. Data are from Quadrini (2020) and the International Centre for Prison Studies (2017-2018).

Table 2.2: Aggregate Crime Rates and Conversions

Crime Type (g)	(*raw* C_g)/C_G	(*raw* C_g)
F		
Robbery	1/24.34	164.99
Burglary	1/132.4	897.91
Larceny	1/351.6	2383.8
Motor Vehicle Theft	1/57.05	386.78
E		
Murder (G)	1	6.7796
Rape	1/5.099	34.571
Aggravated Assault	1/47.77	323.85

The right column (*raw* C_g) provides the grand average rate of each type of crime (g) from 1980-2016 over the whole nation, expressed in incidents per 100,000 people per year. The central column provides the multipliers that convert any raw rates for each type of crime into expression in Standardized Incidents. In other words, a Standardized Incident is 1 Murder, 57 Motor Vehicle Thefts, or so on. These conversions are meant to adjust for the severity of each type of crime - resting on the assumption that more severe crimes occur proportionally less frequently in aggregate than milder crimes do. Crime type (g $= F$) is the average for Financial-Gain crimes (Robbery, Burglary, Larceny, and Motor Vehicle Theft), while (g $=$ E) is the average for Emotional-Gain crimes (Murder, Rape, and Aggravated Assault). Original crime rates must be converted into Standardized Incidents before averages F and E are computed.

Table 2.3: Descriptive Statistics: Crime

Variable	Statistic	1980	1990	2000	2010	2016
$C_{F,\mathrm{t,z}}$	*mean*	5.663	7.734	5.989	4.666	2.904
	stdev	3.467	3.747	2.436	1.586	1.460
	p10	2.400	3.820	2.957	2.541	1.418
	p50	4.641	6.786	5.792	4.621	2.569
	p90	10.29	12.23	8.925	6.642	4.980
$C_{E,\mathrm{t,z}}$	*mean*	5.554	7.693	5.995	5.155	5.849
	stdev	3.033	3.687	2.289	1.671	5.230
	p10	2.257	3.182	3.186	3.060	3.006
	p50	4.874	7.822	5.962	4.995	5.385
	p90	9.567	12.74	9.200	7.452	9.000

$C_{F,\mathrm{z}}$ is the rate of Financial-Gain crimes (Robbery, Burglary, Larceny, and Motor Vehicle Theft), and $C_{E,\mathrm{z}}$ is the rate of Emotional-Gain crimes (Murder, Rape, and Aggravated Assault), in each Commuting Zone, z. Crime rates are expressed in Standardized Incidents per 100,000 people per year. A Standardized Incident is 1 Murder, 24 Robberies, 57 Motor Vehicle Thefts, or so on - see Table 2.2. t restricts by year. Included are only the 250 largest Commuting Zones by population.

Table 2.4: Descriptive Statistics: Earnings and Demographics

Variable	1980	1990	2000	2010	2016
$M_{t,z}$ ($\$^{-1}10^{-4}$)	3.803	2.057	1.251	1.053	0.780
	(0.784)	(0.351)	(0.181)	(0.129)	(0.086)
$Y_{t,z}$ ($\$10^{3}$)	8.551	14.51	21.16	27.96	33.63
	(1.075)	(1.881)	(2.743)	(3.075)	(3.599)
$Q_{t,z}$ (mi^{-2})	85.86	91.36	104.1	110.4	116.2
	(188.2)	(245.4)	(275.9)	(248.1)	(299.8)
$P_{t,z}^{Black}$ (%)	7.529	7.538	8.087	8.657	8.975
	(9.110)	(9.188)	(9.813)	(10.30)	(10.23)
$P_{t,z}^{White}$ (%)	87.81	87.29	83.71	83.45	82.20
	(14.67)	(9.411)	(10.31)	(10.47)	(10.47)
$P_{t,z}^{Young}$ (%)	58.89	58.70	51.76	47.42	48.83
	(8.615)	(3.000)	(3.339)	(3.642)	(3.276)
$P_{t,z}^{Jailed}$ (%)	0.758	1.170	1.958	2.109	2.141
	(0.586)	(0.990)	(1.758)	(1.990)	(1.736)
$P_{t,z}^{College}$ (%)	5.316	15.01	17.84	20.22	21.75
	(1.953)	(3.989)	(4.724)	(5.249)	(5.767)
$P_{t,z}^{Married}$ (%)	67.01	64.82	61.56	53.82	50.57
	(10.08)	(4.452)	(3.953)	(4.321)	(3.962)
$P_{t,z}^{Male}$ (%)	48.07	49.33	50.04	50.52	50.56
	(6.588)	(1.135)	(1.278)	(1.241)	(1.262)

Listed are (unweighted) means of CZ means, not national means. For example, the means for the fraction *White* are higher than the national fraction of Whites would be, because Whites are relatively concentrated in lower population CZs. Standard deviations (of the CZ means) are in parentheses. Y_z is average post-tax earnings, and M_z is average post-tax inverse earnings, per year amongst age 18-64 people in the labor force in each Commuting Zone, z. Q_z is population density. Demographic fractions P are restricted to age 18-64 people who are (except in the case of *Jailed*) not institutionalized. *Young* indicates age 18-40. *Jailed* means currently incarcerated. *College* means at least four years of higher education completed. t restricts by year.

Table 2.5: Highest and Lowest CZs in 2010 Crime Rates

	$C_{F,\mathrm{t,z}}$	CZ	State	$C_{E,\mathrm{t,z}}$	CZ	State
max	9.280	Memphis	TN	10.72	New Orleans	LA
2	9.119	Columbus	GA	10.52	Memphis	TN
3	8.462	Little Rock	AR	9.180	Little Rock	AR
4	8.268	Fayetteville	NC	8.900	Springfield	IL
5	8.093	Florence	SC	8.792	Redding	CA
6	8.028	Houston	TX	8.422	Albuquerque	NM
7	7.727	Jackson	MS	8.359	Detroit	MI
8	7.717	Augusta	GA	8.358	Columbia	SC
9	7.689	Columbus	OH	8.336	Baton Rouge	LA
10	7.684	Bakersfield	CA	8.245	Shreveport	LA
11	7.634	Indianapolis	IN	8.242	Tallahassee	FL
12	7.366	Yakima	WA	8.165	Corpus Christi	TX
13	7.281	San Antonio	TX	8.103	Baltimore	MD
14	7.201	San Francisco	CA	8.103	Florence	SC
15	7.137	Fresno	CA	8.049	Lubbock	TX
60	5.901	Orlando	FL	6.291	Macon	GA
61	5.890	Chicago	IL	6.162	Lansing	MI
p75	5.882	Toledo	OH	6.151	Ocala	FL
63	5.870	Jacksonville	FL	6.141	Fayetteville	NC
64	5.860	Cincinnati	OH	6.137	Beaumont	TX
123	4.655	Killeen	TX	5.038	Athens	GA
124	4.655	Hickory	NC	5.037	Pueblo	CO
p50	4.644	Colorado Springs	CO	5.008	Johnson City	TN
126	4.598	San Jose	CA	4.981	Columbus	GA
127	4.577	Jolpin	MO	4.971	Phoenix	AZ
185	3.445	Columbia	MO	4.075	Boise	ID
186	3.407	Cedar Rapids	IA	4.055	Gastonia	NC
p25	3.399	Albany	NY	4.037	Portland	OR
188	3.379	Lima	OH	4.025	St. Cloud	MN
189	3.377	Kennewick	WA	4.012	San Jose	CA
246	2.016	Eau Claire	WI	2.141	Findlay	OH
247	1.888	State College	PA	1.906	Bangor	ME
248	1.881	Parkersburg	WV	1.795	Southern NJ	NJ
249	1.750	Harrisonburg	MT	1.774	Pikeville	KY
min	1.282	Pikeville	KY	1.748	Wausau	WI

$C_{F,\mathrm{z}}$ is the rate of Financial-Gain crimes (Robbery, Burglary, Larceny, and Motor Vehicle Theft), and $C_{E,\mathrm{z}}$ is the rate of Emotional-Gain crimes (Murder, Rape, and Aggravated Assault), in each Commuting Zone (CZ), z. Crime rates are expressed in Standardized Incidents per 100,000 people per year. A Standardized Incident is 1 Murder, 24 Robberies, 57 Motor Vehicle Thefts, or so on - see Table 2.2. t restricts by year (all 2010 here). Included are only the 250 largest Commuting Zones by population.

Table 2.6: Highest and Lowest CZs in 2010 Earnings and Inverse Earnings

	$Y_{t,z}$	CZ	State	$M_{t,z}$ (10^{-4})	CZ	State
max	\$44,336	Washington	DC	$\$^{-1}1.434$	Redding	CA
2	\$40,636	Northern NJ	NJ	$\$^{-1}1.362$	Medford	OR
3	\$40,384	Southern NJ	NJ	$\$^{-1}1.333$	Chico	CA
4	\$40,189	San Jose	CA	$\$^{-1}1.331$	Gallup	NM
5	\$39,953	San Francisco	CA	$\$^{-1}1.302$	Ocala	FL
6	\$39,216	Boston	MA	$\$^{-1}1.261$	Santa Rosa	CA
7	\$38,881	Bridgeport	CT	$\$^{-1}1.256$	Modesto	CA
8	\$38,715	Baltimore	MD	$\$^{-1}1.253$	Jackson	TN
9	\$38,686	New York City	NY	$\$^{-1}1.222$	Sumter	SC
10	\$37,862	Fredericksburg	VA	$\$^{-1}1.220$	Gastonia	NC
11	\$37,055	Minneapolis	MN	$\$^{-1}1.212$	Hot Springs	AR
12	\$36,469	Philadelphia	PA	$\$^{-1}1.206$	Bakersfield	CA
13	\$36,416	Seattle	WA	$\$^{-1}1.206$	Brownsville	TX
14	\$35,956	Denver	CO	$\$^{-1}1.193$	Pueblo	CO
15	\$35,735	Chicago	IL	$\$^{-1}1.192$	Santa Barbara	CA
60	\$31,582	Davenport	IA	$\$^{-1}1.072$	Rome	GA
61	\$31,581	Cleveland	OH	$\$^{-1}1.072$	Elkhart	IN
p75	\$31,556	Detroit	MI	$\$^{-1}1.071$	Jonesboro	AR
63	\$31,537	Little Rock	AR	$\$^{-1}1.070$	Tuscon	AZ
64	\$31,505	Peoria	IL	$\$^{-1}1.070$	Texarkana	TX
123	\$29,608	Fargo	ND	$\$^{-1}1.003$	Lexington	KY
124	\$29,597	Scranton	PA	$\$^{-1}0.999$	Las Vegas	NV
p50	\$29,587	Kalamazoo	MI	$\$^{-1}0.998$	Columbia	SC
126	\$29,504	Bellingham	WA	$\$^{-1}0.993$	Fort Smith	AR
127	\$29,335	Topeka	KS	$\$^{-1}0.992$	Charleston	SC
185	\$27,704	Paris	TX	$\$^{-1}0.935$	Jackson	MS
186	\$27,675	Tuscaloosa	AL	$\$^{-1}0.934$	Cincinatti	OH
p25	\$27,669	Corpus Christi	TX	$\$^{-1}0.933$	Austin	TX
188	\$27,667	Asheville	NC	$\$^{-1}0.930$	Racine	WI
189	\$27,660	Bloomington	IN	$\$^{-1}0.925$	Dallas	TX
246	\$24,797	Valdosta	GA	$\$^{-1}0.819$	Sioux Falls	SD
247	\$24,340	Sumter	SC	$\$^{-1}0.816$	Charleston	WV
248	\$23,921	Brownsville	TX	$\$^{-1}0.806$	Fredericksburg	VA
249	\$23,690	Hot Springs	AR	$\$^{-1}0.770$	Washington	DC
min	\$23,403	Laredo	TX	$\$^{-1}0.766$	Killeen	TX

Y_z is average post-tax earnings, and M_z is average post-tax inverse earnings, amongst age 18-64 people in the labor force in each Commuting Zone (CZ), z. t restricts by year (all 2010 here). Included are only the 250 largest Commuting Zones by population.

Table 2.7: Elasticities of the Crime Rate

		OLS		2SLS	
Parameter	Regressor	$g = E$	$g = F$	$g = E$	$g = F$
$\Theta_g = \vartheta_g/(1+\vartheta_g)$	$m_{t,z}$	0.355***	0.336***	**0.720*****	**0.783*****
		(0.038)	(0.035)	(0.080)	(0.075)
$\Theta_g\beta_g^y$	$y_{t,z}$	-0.008	0.137***	**0.069**	**0.383*****
		(0.044)	(0.041)	(0.072)	(0.067)
$\Theta_g\beta_g^q$	$q_{t,z}$	0.027***	0.049***	0.036***	0.048***
		(0.004)	(0.004)	(0.005)	(0.004)
$\Theta_g\beta_g^{Black}$	$p_{Black,t,z}$	0.100***	0.078***	0.110***	0.088***
		(0.006)	(0.006)	(0.007)	(0.006)
$\Theta_g\beta_g^{White}$	$p_{White,t,z}$	-0.098***	-0.079***	-0.023	0.019
		(0.029)	(0.027)	(0.033)	(0.031)
$\Theta_g = \vartheta_g/(1+\vartheta_g)$	$m_{t,z}$	0.391***	0.365***	**0.750*****	**0.788*****
		(0.039)	(0.037)	(0.079)	(0.075)
$\Theta_g\beta_g^y$	$y_{t,z}$	0.376***	0.432***	**0.394*****	**0.934*****
		(0.055)	(0.052)	(0.100)	(0.095)
$\Theta_g\beta_g^q$	$q_{t,z}$	0.026***	0.049***	0.028***	0.035***
		(0.004)	(0.004)	(0.005)	(0.004)
$\Theta_g\beta_g^{Black}$	$p_{Black,t,z}$	0.087***	0.065***	0.106***	0.077***
		(0.007)	(0.007)	(0.008)	(0.008)
$\Theta_g\beta_g^{White}$	$p_{White,t,z}$	0.076*	0.045	0.175***	0.255***
		(0.041)	(0.039)	(0.051)	(0.048)
$\Theta_g\beta_g^{Young}$	$p_{Young,t,z}$	-0.026	0.009	0.025	0.222***
		(0.065)	(0.062)	(0.073)	(0.069)
$\Theta_g\beta_g^{Jailed}$	$p_{Jailed,t,z}$	-0.019***	-0.009	-0.014**	-0.007
		(0.006)	(0.006)	(0.006)	(0.006)
$\Theta_g\beta_g^{College}$	$p_{College,t,z}$	-0.225***	-0.169***	-0.131***	-0.221***
		(0.021)	(0.020)	(0.031)	(0.030)
$\Theta_g\beta_g^{Married}$	$p_{Married,t,z}$	-0.721***	-0.553***	-0.577***	-0.646***
		(0.061)	(0.058)	(0.072)	(0.068)
$\Theta_g\beta_g^{Male}$	$p_{Male,t,z}$	0.697***	0.530***	0.326**	0.250**
		(0.115)	(0.110)	(0.131)	(0.124)
Observations		3,610	3,610	3,610	3,610
Weight		$N_{t,z}/N_t$	$N_{t,z}/N_t$	$N_{t,z}/N_t$	$N_{t,z}/N_t$

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the log crime rate $C_{g,t,z}$. Observations are 722 Commuting Zones z, over 5 time periods t corresponding to {1980, 1990, 2000, 2010, 2016}. All columns include year by region effects. N is population count. y is log average earnings. m is log average inverse earnings. q is log population density. p is log fraction of the local population. m, y, and p are restricted to non-institutionalized people of age [18, 64]; m and y are restricted further to those in the labor force. The 2SLS columns instrument for m and y with their Shift-Share counterparts. See previous tables for further definitions.

Table 2.8: Implied Parameter Values

	Baseline		Fully Controlled	
Parameter	$\mathrm{g} = E$	$\mathrm{g} = F$	$\mathrm{g} = E$	$\mathrm{g} = F$
Θ_g	0.720	0.783	0.750	0.788
$\Theta_\mathrm{g}\beta_\mathrm{g}^y$	0	0.383	0.394	0.934
$\beta_\mathrm{g}^y = \Theta_\mathrm{g}\beta_\mathrm{g}^y/\Theta_\mathrm{g}$	0	0.489	0.525	1.185
$\vartheta_\mathrm{g} = \Theta_\mathrm{g}/(1-\Theta_\mathrm{g})$	2.571	3.608	3.000	3.717

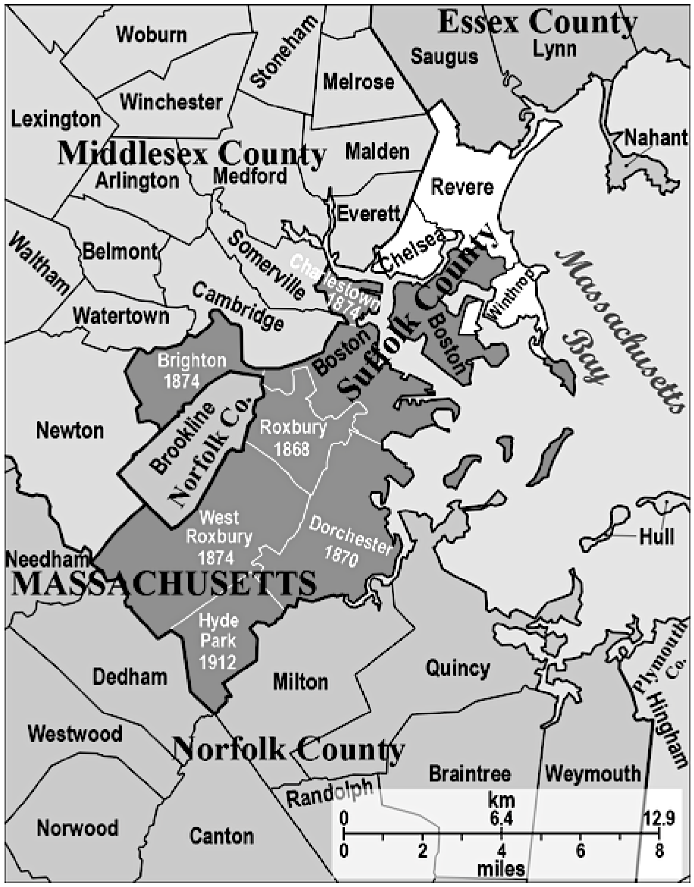
The first two rows are estimates taken from Table 2.7. The two lower rows calculate values of underlying structural parameters implied by the estimates in the first two rows.

Table 2.9: Counterfactual Responses

	$\check{\check{\vartheta}}_\mathrm{g} = -0.2$		$\check{\check{\vartheta}}_\mathrm{g} = +0.2$	
Response	$\mathrm{g} = E$	$\mathrm{g} = F$	$\mathrm{g} = E$	$\mathrm{g} = F$
$\check{\check{\mathrm{K}}}_{\mathrm{g,z}}$ *p90*	+0.091	+0.064	-0.092	-0.079
$\check{\check{\mathrm{K}}}_{\mathrm{g,z}}$ *p50*	+0.112	+0.092	-0.076	-0.062
$\check{\check{\mathrm{K}}}_{\mathrm{g,z}}$ *p10*	+0.140	+0.119	-0.062	-0.044
$\check{\check{\mathrm{C}}}_{\mathrm{g,z}}$ *p90*	-0.123	-0.107	+0.066	+0.046
$\check{\check{\mathrm{C}}}_{\mathrm{g,z}}$ *p50*	-0.101	-0.084	+0.082	+0.066
$\check{\check{\mathrm{C}}}_{\mathrm{g,z}}$ *p10*	-0.084	-0.060	+0.102	+0.085

$\check{\check{\vartheta}}_\mathrm{g}$ is a counterfactual policy change. $\check{\check{\vartheta}}_\mathrm{g} = -0.2$ would indicate that the policy change of choice is to decrease ϑ_g by 20%. The K and C responses give the resulting changes in expected punishment per crime and the crime rate, by Commuting Zone z. F indicates Financial-Gain crimes (Robbery, Burglary, Larceny, and Motor Vehicle Theft), and E indicates Emotional-Gain crimes (Murder, Rape, and Aggravated Assault). ϑ represents the government's distaste for criminal punishment. A lower value of ϑ coincides with a tougher stance on crime.

Figure 2: County Borders around Boston, MA



The map of the Boston, MA area illustrates why smaller geographical units such as cities or counties may be inappropriate to capture local crime markets. People frequently commute between Somerville, Cambridge, Boston (Basu and Ferreira, 2021; Florida and Mellander, 2016; Matarazzo et al., 2017), and Brookline. These four cities belong to three distinct counties - Middlesex, Suffolk, and Norfolk. Commuting Zones, defined based on actual commuting patterns in 1990, are meant to account for the vast majority of such local spillovers. The Boston CZ (see Figure 3) includes all of the counties pictured.

Figure 3: Boston, MA Commuting Zone

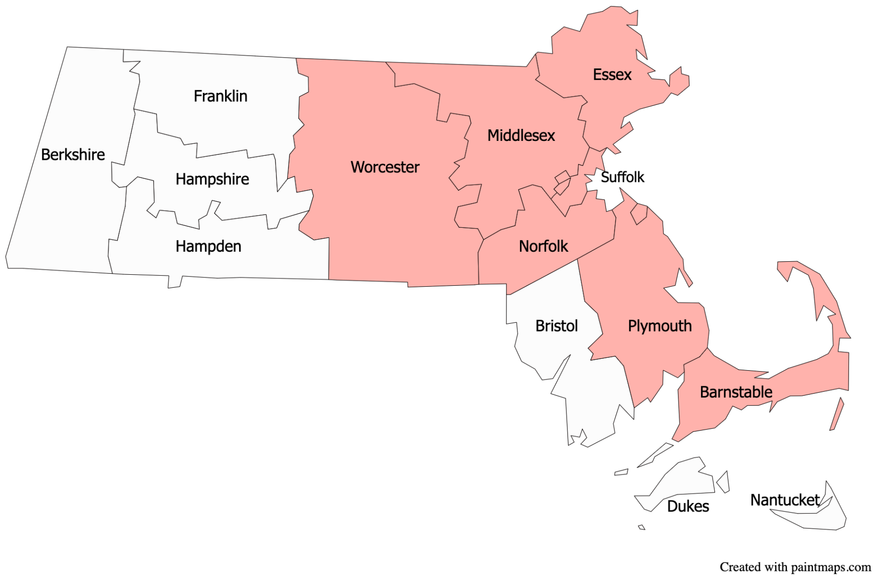
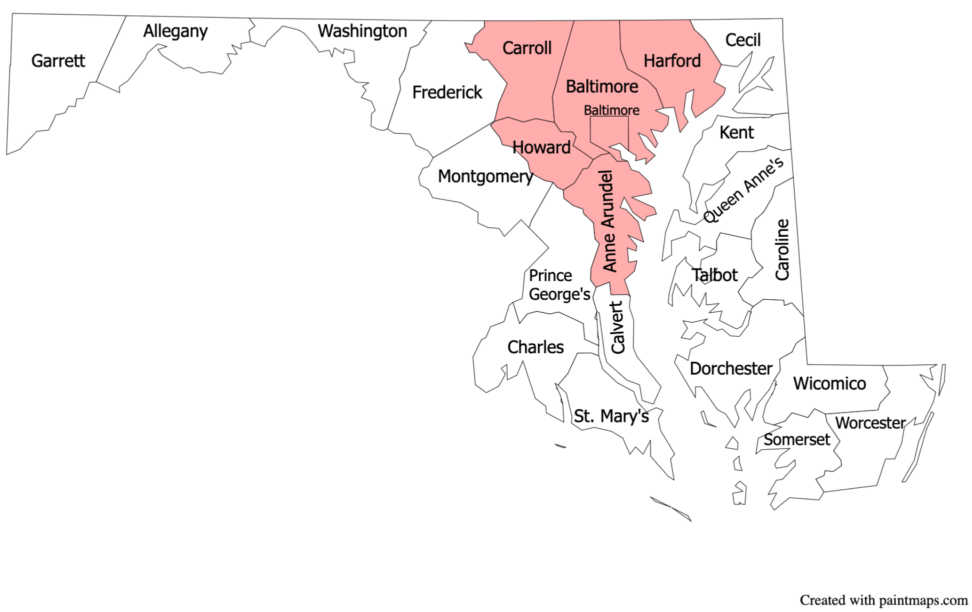


Figure 4: Baltimore, MD Commuting Zone



Most CZs are contained within state lines, but not all. The Washington, DC CZ contains Maryland counties from Frederick to St Mary's (above), as well as several Virginia counties.

2.13 Appendix

2.13.1 Data Sources

As discussed in Section 2.3, we use county level summaries of the Decennial Census and American Community Survey (ACS), and also IPUMS USA microsamples of the same, in concert. We take county level summaries from Social Explorer,[92] particularly, 1970, 1980, 1990, and 2000 US Decennial Census summaries, and ACS 2008-2012 and 2015-2019 5-Year Estimate summaries. From IPUMS USA, we take the 1970 1% metro fm2, 1980 1% metro, 1990 1% unwt, 2000 1% unwt, 2010 ACS 1%, and 2019 ACS 1% samples. We use Dorn's crosswalk files[93] to integrate all of the above data into a panel of 722 Commuting Zones (CZs) by 6 decadal time periods.

At the CZ level, we compute a separate IPUMS measured and Social Explorer (SE) measured version of each variable, as available. In IPUMS, the source observations are of individual persons. Each variable we compute is a weighted average per CZ-year cell, weighted by perwt (person weight) and afactor, as described in Dorn (2009). In SE, the source observations are of counties, which group directly into CZs. For most SE measured variables, we take a population weighted average of the counties in each CZ-year cell.

For every variable at the CZ level, both in IPUMS and in SE, we also calculate versions of the same at two higher levels of aggregation, which we call Sub Region and Super Region. Sub Regions are the 9 regional divisions used by the Census,[94] but split into Urban and Rural CZs within each, for $9 \times 2 = 18$ in total. We define an Urban CZ as one in which, in the 1990 IPUMS sample, there were at least 80,000 people identified as living in an urban area as defined by the Census. Super Regions

[92] https://www.socialexplorer.com/explore-tables

[93] See Autor and Dorn (2013) and Autor, Dorn and Hanson (2019).

[94] The divisions are New England (CT, MA, ME, NH, RI, VT), Mid-Atlantic (NJ, NY, PA), East North Central (IL, IN, MI, OH, WI), West North Central (IA, KS, MN, MO, NE, ND, SD), South Atlantic (DC, DE, FL, GA, MD, NC, SC, VA, WV), East South Central (AL, KY, MS, TN), West South Central (AR, LA, OK, TX), Mountain (AZ, CO, ID, MT, NV, NM, UT, WY), and Pacific (CA, OR, WA).

are the 4 main regions used by the Census, also split into Urban and Rural CZs within each, for 8 in total. The main regions are Northeast (New England, Mid-Atlantic), Midwest (East North Central, West North Central), South (South Atlantic, East South Central, West South Central), and West (Mountain, Pacific).

2.13.2 Data Transformations

We construct final versions of each variable at the CZ level based on weighted averages of several versions of the variable. The weights given to each version are to some extent increasing in the number of individual level observations underlying each. They include IPUMS sourced CZ level, Sub Region level, and Super Region level versions, and transformations of a closely related SE sourced variable at the CZ level, transformed at the Sub Region and Super Region levels. Transformations are of the form,

$$\mathrm{v_z} = \frac{\mathrm{v}_{\mathrm{Z}}^{IPUMS}}{\mathrm{V}_{\mathrm{Z}}^{IPUMS}} \mathrm{V}_{\mathrm{z}}^{SE}$$

where V is a variable closely related to v that is available through SE, z is CZ, and Z is Region or Super Region. For example, with v as the average earnings of non-institutionalized people of age 18-64, V may be the average earnings of all people above age 15.

The use of these weighted averages serves three purposes. First, it guarantees a unique value to every CZ-year cell, for every variable. Second, it uses all available information in a consistent manner, distributing weight in accordance with the amount of information underlying each source estimate, which should improve overall accuracy. Third, the inclusion of Sub Region and Super Region level estimates within each final CZ level estimate accounts for some amount of spillover amongst neighboring CZs, albeit of an arbitrary magnitude. For the average CZ, the IPUMS CZ level estimate accounts for about 80% of the final estimate.

Our transformation technique, respective to any given case of v and V, hinges on the assumption that the ratio of v to V in z is equal to that in Z. v and V must be closely related for this assumption to be plausible. A concrete example may be: if age 16+ people in rural New Hampshire earn 4% more than the average for age 16+ people in rural New England, then age 18-64 people in rural New Hampshire are assumed to earn more 4% more than the average for age 18-64 people in rural New England. Symbolically,

$$\mathrm{v_z} = \delta \mathrm{v_Z} \ , \ \mathrm{V_z} = \delta \mathrm{V_Z}$$

where δ is a scalar, such as 1.04. Therefore,

$$\frac{\mathrm{v_z}}{\mathrm{V_z}} = \frac{\mathrm{v_Z}}{\mathrm{V_Z}}$$

We then assume that SE accurately measures $\mathrm{V_z}$, and IPUMS accurately measures $\mathrm{v_Z}/\mathrm{V_Z}$.

2.13.3 Tax

As referenced in the model, we adjust earnings earnings by federal tax, redistributing half of the tax revenue as an in-kind earnings baseline $\underline{\mathrm{Y}}_\mathrm{t}$. The Census and ACS, via IPUMS USA, provide pre-tax earnings, $\mathrm{Y_i}$, per individual person in the sample, i. National average pre-tax earnings are $\bar{\mathrm{Y}}_\mathrm{t} = \mathrm{N_t^{-1}} \sum_\mathrm{i} \mathrm{Y_i}$, where N is population count, and t is time period. To arrive at post-tax earnings, we apply the tax function of Heathcote, Storesletten and Violante (2017),

$$\text{post-tax } \mathrm{Y_i} \propto (\text{pre-tax } \mathrm{Y_i})^{1-\tau} \ , \ \tau = 0.181 \tag{40}$$

Heathcote et al. show that this function is a close match to the actual federal + state tax and transfer schedule in the US. That is, although the function is fit ($\tau = 0.181$) to capture state taxes (on average) in addition to federal tax, the function itself does not vary by state, so we refer to it throughout the paper simply as federal tax.

In conjunction with (40), it is necessary also to assume the aggregate amount of tax that is collected. We assume total tax is 20% of total pre-tax earnings, that is,

$$N_t^{-1} \sum_i (Y_i - \tilde{\tau} Y_i^{1-\tau}) = 20\% \; \bar{Y}_t \tag{41}$$

$Y_i - \tilde{\tau} Y_i^{1-\tau}$ being the net tax that is collected per individual i, and Y_i being pre-tax earnings. (41) enables algebraic derivation of the constant of proportionality $\tilde{\tau}$, which resolves the exact value of each individual's tax payment $(Y_i - \tilde{\tau} Y_i^{1-\tau})$.

$$\tilde{\tau} = \frac{(1 - 0.2) \sum_i Y_i}{\sum_i Y_i^{1-0.181}} \tag{42}$$

Because $\tau = 0.181 > 0$, this is a progressive tax, taking more in percentage terms from individuals with higher pre-tax earnings Y_i.

Despite being a single function that is applied to all individuals i equally regardless of their location z, the tax (40) does affect the post-tax earnings averages Y_z across locations z differently. Table 2.6, which ranks CZs by their post-tax earnings averages Y_z, would not only list different values if not for the tax, but would be in a different order. For example, San Francisco may have a higher pre-tax average than Southern NJ, but with a relatively more unequal distribution. Because the tax is progressive, it takes a larger share from very high earning individuals, so may take substantially more in tax from San Francisco than it takes from Southern NJ. This could result in San Francisco's post-tax average falling slightly below Southern NJ's, despite the order being the other way around in pre-tax averages.

We assume that, in addition to paying the tax (40), people benefit from the tax revenue. Particularly, we assume that half of the tax budget is redistributed equally as in-kind benefits $\underline{Y}_t$. We add $\underline{Y}_t$ to each individual's post-tax earnings before calculating the averages $Y_{t,z}$ and $M_{t,z}$. Because the tax budget is 20% of total pre-tax earnings, $\underline{Y}_t$ is 10%, that is, $\underline{Y}_t = 10\% \; \bar{Y}_t$. This, in addition to the progressivity $(\tau > 0)$ of the tax payment schedule, further moderates (reduces inequality in) the earnings distribution, both within locations z and across. That is, despite being paid out exactly equally across individuals, $\underline{Y}_t$ makes a larger difference in relative terms for people with lower earnings.

2.13.4 First and Second Order Conditions

This section shows that the equilibrium crime rate and expected level of punishment per crime given in Section 2.5 optimize the offender's utility function and government's loss function, respectively. Subscripts g (for crime type) and t (for time period) are suppressed for readability.

The offender i's utility function weighs his expected personal benefit per crime B_z against his expected punishment per crime K_z, in each Commuting Zone z.

$$\begin{aligned} U(C_{i,z}) &= \frac{B_z}{1-\alpha}(C_{i,z})^{1-\alpha} - \frac{K_z}{2-\alpha}(C_{i,z})^{2-\alpha}(Y_{i,z}+\underline{Y})\ ,\ \alpha \in [0,1] \\ U'(C_{i,z}) &= B_z(C_{i,z})^{-\alpha} - K_z(C_{i,z})^{1-\alpha}(Y_{i,z}+\underline{Y}) \\ U''(C_{i,z}) &= -\alpha B_z(C_{i,z})^{-1-\alpha} - (1-\alpha)K_z(C_{i,z})^{-\alpha}(Y_{i,z}+\underline{Y}) \end{aligned} \tag{43}$$

The first order condition is:

$$U'(C_{i,z})\Bigg|\left\{C_{i,z} = C^*_{i,z}\right\} = 0$$

$$\Leftrightarrow\ B(C^*_{i,z})^{-\alpha} = K(C^*_{i,z})^{1-\alpha}(Y_{i,z}+\underline{Y})\ \Leftrightarrow\ C^*_{i,z} = (Y_{i,z}+\underline{Y})^{-1}\frac{B_z}{K_z}$$

The second order condition,

$$U''(C_{i,z})\Bigg|\left\{C_{i,z} = C^*_{i,z}\right\} < 0$$

follows straightforwardly from (43), regardless of $C^*_{i,z}$. Because $\alpha \in [0,1]$, at least one of $-\alpha$ and $-(1-\alpha)$ must be negative, and neither can be positive. C, B, K are all positive by construction. Therefore only negative terms are added together.

The government's loss function weighs the direct social cost of punishment K_z itself, against the social cost of crime, which responds to punishment indirectly, $C(K_z)$.

$$\begin{aligned}
L(K_z) &= \vartheta^{-1}K_z^{\vartheta} + C(K_z) = \vartheta^{-1}K_z^{\vartheta} + \frac{M_zB_z}{K_z} \\
L'(K_z) &= K_z^{\vartheta-1} - \frac{M_zB_z}{K_z^2} \\
L''(K_z) &= (\vartheta-1)K_z^{\vartheta-2} + 2\frac{M_zB_z}{K_z^3}
\end{aligned} \tag{44}$$

The first order condition is:

$$L'(K_z)\bigg|\left\{K_z = K_z^*\right\} = 0$$

$$\Leftrightarrow\ (K_z^*)^{\vartheta-1} = \frac{M_zB_z}{(K_z^*)^2}\ \Leftrightarrow\ (K_z^*)^{\vartheta-1+2} = M_zB_z\ \Leftrightarrow\ K_z^* = (M_zB_z)^{1/(1+\vartheta)}$$

Because $L(K_z)$ is a loss function, it is optimized at a minimum, where the second derivative is positive. Thus the second order condition is:

$$L''(K_z)\bigg|\left\{K_z = K_z^* = (M_zB_z)^{1/(1+\vartheta)}\right\} > 0$$

The derived K_z^* can be plugged in to (44), yielding:

$$L''(K_z)\bigg|\left\{K_z = K_z^*\right\} = (\vartheta-1)(M_zB_z)^{(\vartheta-2)/(1+\vartheta)} + 2(M_zB_z)^{1-3/(1+\vartheta)}$$

Notice that $1-3/(1+\vartheta) = (1+\vartheta-3)/(1+\vartheta) = (\vartheta-2)/(1+\vartheta)$. Therefore the above simplifies to:

$$(\vartheta-1+2)(M_zB_z)^{(\vartheta-2)/(1+\vartheta)} = (1+\vartheta)(M_zB_z)^{(\vartheta-2)/(1+\vartheta)} > 0$$

$(1+\vartheta)$ must be positive because $\vartheta > 0$, and M_zB_z are positive by construction.

2.13.5 Counterfactual Responses

The main step in deriving counterfactual responses to changes in ϑ involves inverting the equilibrium equations (34) and (35). With (crime type) g and (time) t subscripts suppressed, these equilibrium equations are:

$$\begin{aligned}\mathrm{K_z} &= (\mathrm{M_zB_z})^{1/(1+\vartheta)} = (\mathrm{C_z})^{1/\vartheta} \\ \mathrm{C_z} &= (\mathrm{M_zB_z})^{\vartheta/(1+\vartheta)} = (\mathrm{K_z})^{\vartheta}\end{aligned}$$

Inverted versions of the same are:

$$\begin{aligned}\mathrm{M_zB_z} &= (\mathrm{K_z})^{1+\vartheta} \\ \mathrm{M_zB_z} &= (\mathrm{C_z})^{(1+\vartheta)/\vartheta}\end{aligned} \tag{45}$$

Because nothing in $\mathrm{M_z}$ or $\mathrm{B_z}$ is a function of ϑ, counterfactual equilibrium values $\check{\mathrm{K}}_\mathrm{z}$ and $\check{\mathrm{C}}_\mathrm{z}$ follow from the counterfactual $\check{\vartheta}$ of choice as:

$$\begin{aligned}\check{\mathrm{K}}_\mathrm{z} &= (\mathrm{M_zB_z})^{1/(1+\check{\vartheta})} \\ \check{\mathrm{C}}_\mathrm{z} &= (\mathrm{M_zB_z})^{\check{\vartheta}/(1+\check{\vartheta})}\end{aligned}$$

Although $\mathrm{M_zB_z}$ is not directly observed, it can be replaced using the inverted factual equilibrium (45):

$$\begin{aligned}\check{\mathrm{K}}_\mathrm{z} &= ((\mathrm{K_z})^{1+\vartheta})^{1/(1+\check{\vartheta})} = (\mathrm{K_z})^{(1+\vartheta)/(1+\check{\vartheta})} \\ \check{\mathrm{C}}_\mathrm{z} &= ((\mathrm{C_z})^{(1+\vartheta)/\vartheta})^{\check{\vartheta}/(1+\check{\vartheta})} = (\mathrm{C_z})^{(\check{\vartheta}+\vartheta\check{\vartheta})/(\vartheta+\vartheta\check{\vartheta})}\end{aligned} \tag{46}$$

Of primary interest are how each variable changes in the counterfactual relative to its factual counterpart:

$$\check{\check{\vartheta}} := (\check{\vartheta} - \vartheta)/\vartheta \ , \ \check{\check{\mathrm{K}}}_\mathrm{z} := (\check{\mathrm{K}}_\mathrm{z} - \mathrm{K_z})/\mathrm{K_z} \ , \ \check{\check{\mathrm{C}}}_\mathrm{z} := (\check{\mathrm{C}}_\mathrm{z} - \mathrm{C_z})/\mathrm{C_z} \tag{47}$$

For example, $\check{\check{\vartheta}} = 0.1$ would mean that the counterfactual policy choice is to increase

ϑ by 10%. Combining (46) and (47):

$$\check{\check{K}}_z = ((K_z)^{(1+\vartheta)/(1+\check{\vartheta})} - K_z)/K_z = (K_z)^{(1+\vartheta)/(1+\check{\vartheta})-1} - 1 = (K_z)^{-\vartheta\check{\check{\vartheta}}/(1+\check{\vartheta})} - 1$$

$$\check{\check{C}}_z = ((C_z)^{(\check{\vartheta}+\vartheta\check{\vartheta})/(\vartheta+\vartheta\check{\vartheta})} - C_z)/C_z = (C_z)^{(\check{\vartheta}+\vartheta\check{\vartheta})/(\vartheta+\vartheta\check{\vartheta})-1} - 1 = (C_z)^{\check{\check{\vartheta}}/(1+\check{\vartheta})} - 1$$

To simplify further - to functions of only the observed C_z, estimated ϑ, and chosen $\check{\check{\vartheta}}$ - apply the facts that $K_z = C_z^{1/\vartheta}$ (equilibrium condition), and $\check{\vartheta} = \vartheta(1+\check{\check{\vartheta}})$ (definition of $\check{\check{\vartheta}}$):

$$\check{\check{K}}_z = (C_z)^{-\check{\check{\vartheta}}/(1+\vartheta(1+\check{\check{\vartheta}}))} - 1$$

$$\check{\check{C}}_z = (C_z)^{\check{\check{\vartheta}}/(1+\vartheta(1+\check{\check{\vartheta}}))} - 1$$

The above are equilibrium formulas that predict how expected punishment per crime K and crime rates C would adjust under a different social loss function. A higher value of ϑ (so $\check{\check{\vartheta}} > 0$) would mean society (the government) placing increased weight on the pain inherent in criminal punishment in and of itself. A lower value of ϑ (so $\check{\check{\vartheta}} < 0$) would mean the opposite - a society that is less sensitive to the pain of punishment (relative to the harm caused by crime), and therefore a tougher stance against crime.

3 Novel Shift-Share Instruments and Their Applications

3.1 Introduction

The Shift-Share approach is a powerful and flexible framework for developing instrumental variables for causal identification. Many papers have surveyed and examined the validity of particular Shift-Share instruments, as given objects in the literature.[95] This paper rather examines the creative process by which one develops or arrives at Shift-Share instruments to begin with. In other words, this paper crystallizes the essential features of Shift-Share instruments, the purpose of each feature, and the scope of how each can be extended.

The upcoming Section 3.2 forms a foundation for the remainder of the paper: It defines the core elements of classical Shift-Share instruments, and develops a unified system of notation into which I will translate disparate papers' models for the purpose of comparison. In following sections, I discuss the history of Shift-Share instruments in the literature, ranging from standard to more exotic variants. By closely comparing many distinct variants, I illustrate both that which is most essential about Shift-Share instruments, and the wealth of different ways in which they may be adapted.

The primary contribution of this paper may be the simple framework it proposes for how to understand Shift-Share instruments. The essence of the Shift-Share approach is to decompose the endogenous explanatory variable as an accounting identity with multiple component parts; preserve that which is most exogenous in the accounting identity, and neutralize that which is most endogenous. Endogeneity is neutralized via delocalizations over space and time. That is, the more endogenous component - the Shift vector - is replaced with nonlocal averages (a strong delocalization); and the more exogenous component - the Share vector - is lagged (a weaker delocalization).

As an additional contribution, I develop several new varieties of Shift-Share instruments, particularly for explanatory variables that are distribution summaries other

[95]See Baum-Snow and Ferreira (2015), Goldsmith-Pinkham, Sorkin and Swift (2020).

than means. That is, I develop general formulas for Shift-Share instruments for variances, skews, mean absolute deviations around central points, and Gini coefficients. These formulas strongly emphasize the importance of deriving the instrument from an accounting identity of the explanatory variable - a central theme of the paper. As an application, I measure the elasticity of single parenting with respect to earnings inequality, using multiple alternative instruments for the Gini coefficient of earnings. Empirically, I find that instruments using only either lags or a Shift Delocalization (but not both) do correct bias, but only part way. This underlines the importance of using both, as in the full Shift-Share instrument.

3.2 Notation and Core Principles

I begin by defining some standardized notation, which although technical in nature, is critical for understanding the topic at hand. This paper considers econometric models from many different papers, each of which has its own completely distinct system of notation. By translating these out of their native notation into a common shared system, I hope to enable far greater insight into their substantive similarities and differences.

Shift-Share instruments are panel data objects. As such, they involve a time dimension, which I denote as t, and cross sectional units, which I denote as z. These cross sectional units z traditionally are geographical units, such as Metropolitan Statistical Areas (MSAs), meant to capture local labor markets. However, they needn't necessarily be geographical units. I call these units z Localities - not in a concrete sense, but rather in an abstract sense of "local" as opposed to "global".

I denote the typical endogenous explanatory variable of interest as $\mathrm{X}_{\mathrm{t,z}}$. For example, $\mathrm{X}_{\mathrm{t,z}}$ may be the average wage in Locality z (such as z = *Boston*) in time period t (such as t = *2010*). Generally, any average is (at least implicitly) an average over individuals, i.

$$\mathrm{X_{t,z}} = \mathrm{N_{t,z}^{-1}} \sum_{\mathrm{i \in t,z}} \mathrm{X_i}$$

where $\mathrm{X_i}$ the value of the variable per individual (such as an individual person's wage), and $\mathrm{N_{t,z}}$ is the count of individuals in Locality z in time period t. The appropriate kind of count $\mathrm{N_{t,z}}$ may depend on the variable $\mathrm{X_i}$. In particular, it is often the case that we should count only employed individuals, such as when $\mathrm{X_i}$ is only defined for those who are employed (wages may be such a case). To emphasize this, I use $\mathrm{NE_{t,z}}$ to denote counts of employed persons per se.

The Shift-Share approach begins with a decomposition over industrial, occupational, or similar categories, which I denote as o. For simplicity, I refer to any such categorization as Industries in an abstract sense. The essential feature of such a categorization - whether industrial, occupational, or otherwise - is that the Shares they define should be less sensitive to endogeneity than $\mathrm{X_{t,z}}$ is. For example, each employed person works in a particular industry, o. Industrial Shares are $\mathrm{NE_{o,t,z}/NE_{t,z}}$,[96] that is, the fraction of the employed who are employed in industry o. It may be that in z = *Detroit* in t = *1990*, there are $\mathrm{NE_{t,z}}$ = 600,000 total workers, and $\mathrm{NE_{o,t,z}}$ = 120,000 workers in o = *Manufacturing.* The Share would then be 120,000/600,000 = 20%. That about 20% of Detroit's employment opportunities are in Manufacturing would be considered less sensitive to endogeneity than (for example) the average wage $\mathrm{X_{t,z}}$ in Detroit is, because the former is a result of geographical and historical factors that cannot be easily adjusted.

A "share" is simply a fraction; therefore other kinds of "shares" may be confused for the Shares of a Shift-Share instrument. A salient case in point is that Industries' shares of Localities are not the same as Localities' shares of Industries. Many authors

[96] Shares may also be $\mathrm{N_{o,t,z}/N_{t,z}}$, if $\mathrm{X_i}$ is defined for the whole population rather than just the employed.

have understood Industries' shares of Localities, that is,

$$\mathrm{NE_{o,t,z}/NE_{t,z}}$$

to be the proper Shares of a Shift-Share instrument. However, Localities' shares of industries have sometimes been used in a similar way.

$$\mathrm{NE_{o,t,z}/NE_{o,t}}$$

The numerator is the same, but the denominator is different. Returning to the example from the previous paragraph, $\mathrm{NE_{o,t,z}} = 120{,}000$ is still the number of Manufacturing workers in Detroit. But now the denominator $\mathrm{N_{o,t}}$ is the national count of Manufacturing workers - let's say 3 million - rather than the local count of all workers ($\mathrm{NE_{t,z}} = 600{,}000$). Although numerically different, Localities' shares of Industries may be considered insensitive to endogenetiy for much the same reasons that Industries' shares of Localities are, and therefore may feature similarly in some instruments.

Shares of Localities - i.e., Industries' shares of Localities, rather than Localities' shares of Industries - possess a unique feature that make them the most natural choice. This is that for *any* variable $\mathrm{X_{t,z}}$ that is an average, the following is true as an accounting identity.

Classical Accounting Identity:

$$\mathrm{X_{t,z}} = \sum_{\mathrm{o}} \mathrm{X_{o,t,z} \cdot N_{o,t,z}/N_{t,z}}$$

To see why this is so, one must simply unpack definitions:

$$\mathrm{X_{t,z}} = \mathrm{N_{t,z}^{-1}} \sum_{\mathrm{i \in t,z}} \mathrm{X_i} \;\Leftrightarrow\; \sum_{\mathrm{i \in t,z}} \mathrm{X_i} = \mathrm{X_{t,z} \cdot N_{t,z}}$$

$$X_{o,t,z} = N_{o,t,z}^{-1} \sum_{i \in o,t,z} X_i \;\Leftrightarrow\; \sum_{i \in o,t,z} X_i = X_{o,t,z} \cdot N_{o,t,z}$$

And pivotally, it is a truism that:[97]

$$\sum_{i \in t,z} X_i = \sum_{o} \sum_{i \in o,t,z} X_i \tag{48}$$

Plugging the definition equations above into (48), we have,

$$X_{t,z} \cdot N_{t,z} = \sum_{o} X_{o,t,z} \cdot N_{o,t,z}$$

And because $N_{t,z}$ is constant with respect to o, it can be rewritten (divide both sides by $N_{t,z}$) inside the summation, as the denominator under $N_{o,t,z}$.

To arrive at a Shift-Share instrument, one alters the Shift-Share accounting identity in order to preserve that which is most exogenous in $X_{t,z}$, and remove that which is most sensitive to endogeneity.

Classical Accounting Identity:

$$X_{t,z} = \sum_{o} \overbrace{X_{o,t,z}}^{\text{Shift}} \cdot \overbrace{N_{o,t,z}/N_{t,z}}^{\text{Share}}$$

The primary alteration is to delocalize the Shift over Localities z. That is, replace the Shift $X_{o,t,z}$ with the Delocalized Shift, $X_{o,t}$.

Delocalized Shift Instrument:

$$\tilde{X}_{t,z} = \sum_{o} \overbrace{X_{o,t}}^{\text{Shift}} \cdot \overbrace{N_{o,t,z}/N_{t,z}}^{\text{Share}}$$

[97] The sum of the whole is equal to the sum of (mutually exclusive and collectively exhaustive) subset sums.

If $X_{o,t,z}$ is the average wage of Manufacturing workers in Detroit, $X_{o,t}$ is the average wage of Manufacturing workers nationally - a *delocalized* average.[98]

The Shift Delocalization is both the strongest and the most distinctive aspect of Shift-Share instruments' claim to exogeneity. As discussed earlier in this section, Industries o should be such that the Shares they define are less sensitive to endogeneity than $X_{t,z}$ is. The Shares' complement - the local Shift vector $X_{o,t,z}$ - is by contrast the part of $X_{t,z}$ *most* sensitive to endogeneity. Replacing $X_{o,t,z}$ with $X_{o,t}$ annihilates this main part of $X_{t,z}$'s endogeneity, because national averages $X_{o,t}$ cannot be a function of anything at the level of isolated Localities z. As an example, suppose we are interested in the effects of wages on hours of work supplied. The average wage of Manufacturing workers in Detroit, $X_{o,t,z}$, is sensitive to reverse causality: an abundance of able workers in Detroit may depress their wages. However, the national average wage of Manufacturing workers, $X_{o,t}$, is not sensitive to the local overabundance of workers in Detroit. It represents rather the portion of $X_{o,t,z}$ that is independent of Detroit's local conditions.

Because the main part of $X_{t,z}$'s endogeneity is in the local Shifts, the Shift Delocalization removes this main part. However, there remains a question of whether some endogeneity still lingers in the Shares. For this reason, a secondary alteration to the accounting identity is made: Lag the Shares, such as to a base period *t*.

Classical Accounting Identity:

$$X_{t,z} = \sum_{o} \overbrace{X_{o,t,z}}^{\text{Shift}} \cdot \overbrace{N_{o,t,z}/N_{t,z}}^{\text{Share}}$$

[98] Alternatively to the national average $X_{o,t}$, some authors have used the average over all Localities excluding z, $X_{o,t,\neg z}$. For large cross sections, these are essentially the same, because each individual Locality plays only a very small role in the national average.

Delocalized Shift Instrument:

$$\tilde{\mathrm{X}}_{\mathrm{t,z}} = \sum_{\mathrm{o}} \overbrace{\mathrm{X}_{\mathrm{o,t}}}^{\text{Shift}} \cdot \overbrace{\mathrm{N}_{\mathrm{o,t,z}}/\mathrm{N}_{\mathrm{t,z}}}^{\text{Share}}$$

Shift-Share Instrument:

$$\tilde{\mathrm{X}}_{\mathrm{t,z}} = \sum_{\mathrm{o}} \overbrace{\mathrm{X}_{\mathrm{o,t}}}^{\text{Shift}} \cdot \overbrace{\mathrm{N}_{\mathrm{o},t,\mathrm{z}}/\mathrm{N}_{t,\mathrm{z}}}^{\text{Share}}$$

The purpose for lagging the Shares is similar to that for delocalizing the Shifts. A lag is a kind of delocalization, only over time rather than over the cross section. However, it is a weaker kind of delocalization, as the lagged Share retains a connection to the Locality z, while Shift does not at all. The mildness of the remedy is apropos to the mildness of the problem. That is, the Shares would already be considered mostly or plausibly exogenous even in time period t; the lag is just an added guarantee.

Alternatively, it may be questioned whether it is even necessary for the Shares to be exogenous. Goldsmith-Pinkham, Sorkin and Swift (2020) discuss that there are two designs by which Shift-Share instruments may be viewed to be exogenous. That is, researchers may view the exogeneity either as "coming from the Shares," or as "coming from the shocks [Shifts]." I would phrase this differently in that, as discussed in the previous paragraphs, the Delocalized Shifts are definitely exogenous. The question rather is whether this is sufficient: Does the product of an exogenous (Delocalized Shift) vector and an endogenous (Share) vector come out as exogenous like the Shift, or as endogenous like the Share? Borusyak, Hull and Jaravel (2022) present a scenario by which the former is true: The product is exogenous by force of the Shifts' exogeneity alone. But Adão, Kolesár and Morales (2019) show this to be tenuous, and as Goldsmith-Pinkham, Sorkin and Swift (2020) also discuss, most researchers have conceived of Shift-Share instruments as relying on both the

(Delocalized) Shifts and the (Lagged) Shares to be exogenous.

In addition to delocalizing the Shifts and lagging the Shares, there may be a zeroth or third step, which is meant to account for effects of unobservables. The Shift Delocalization and Share Lag are powerful measures that make the instrument causally prior with respect to any variable that is observed in the estimation, including the outcome variable. However there is little they can do with respect to unobserved factors that might be correlated with both the instrument and the outcome. For such unobservables, the Shift-Share approach relies on standard methods of time differencing, fixed effects and other controls. As a zeroth step, $\mathrm{X}_{\mathrm{t,z}}$ may be defined as a time difference, such as a growth rate; or as a third step, both the estimating equation and the instrument may be time differenced. This will have the implication of canceling out time-invariant unobservables, though not necessarily time-invariant unobservables. Alternatively as a third step, the instrument may be accompanied by fixed effects or other controls meant absorb the effects of unobservables.

3.3 Classical Variants

Bartik (1991) and the closely related Blanchard and Katz (1992) are broadly credited with introducing Shift-Share instruments to modern economics literature. Both investigate the effects of job growth on wages, employment and unemployment rates as outcomes, in Metropolitan Statistical Areas (MSAs) as labor markets.[99] In other words, they ask how local job growth is absorbed: does it pull locals into employment out from unemployment, or out from non-participation (not looking for a job at all), or does it rather pull migrant workers from elsewhere; and how are local wages impacted. Although these are multiple outcomes of interest, the authors are interested in the effects on each outcome of only one main causal factor, that is, job growth. Therefore a single instrumental variable for job growth is sufficient to identify all the

[99]Job (growth) refers to (change over time in) the absolute number of employed workers - unlike employment or unemployment *rates*, which are divided by a local population count even in level terms.

effects in question. Bartik (1991) develops such an instrument, and Blanchard and Katz (1992) import it directly from the former.

The concept of a Shift-Share decomposition has earlier precedent; but Bartik (1991) may have been the first to apply it to create an instrumental variable.[100] As a general framework, the Shift-Share instrument is essentially an altered version of the endogenous explanatory variable itself. Therefore it is applicable in a great variety of settings, leading to widespread adoption. Nonetheless, there may be ambiguity as to what "Shift-Share" is actually supposed to mean. The below quote from Bartik (1991) may be helpful in clarifying his terminology:

> A shift-share analysis decomposes MSA growth into three components: a national growth component, which calculates what growth would have occurred if all industries in the MSA had grown at the all-industry national average; a share component, which calculates what extra growth would have occurred if each industry in the MSA had grown at that industry's national average; and a shift component, which calculates the extra growth that occurs because industries grow at different rates locally than they do nationally.

In my notation, this can be translated as follows:

$$\mathrm{X_{t,z}} = \sum_{\mathrm{o}} \Big(\overbrace{\mathrm{X_t}}^{\substack{\text{“National}\\ \text{Component”}}} + \overbrace{\mathrm{X_{o,t}} - \mathrm{X_t}}^{\substack{\text{“Share}\\ \text{Component”}}} + \overbrace{\mathrm{X_{o,t,z}} - \mathrm{X_{o,t}}}^{\substack{\text{“Shift}\\ \text{Component”}}} \Big) \cdot \mathrm{NE_{o,t,z}/NE_{t,z}}$$

$$\mathrm{X_{t,z}} = (\mathrm{NE_{t,z}} - \mathrm{NE_{t-1,z}})/\mathrm{NE_{t-1,z}}$$

$$\mathrm{X_t} = (\mathrm{NE_t} - \mathrm{NE_{t-1}})/\mathrm{NE_{t-1}}$$

$$\mathrm{X_{o,t}} = (\mathrm{NE_{o,t}} - \mathrm{NE_{o,t-1}})/\mathrm{NE_{o,t-1}}$$

$$\mathrm{X_{o,t,z}} = (\mathrm{NE_{o,t,z}} - \mathrm{NE_{o,t-1,z}})/\mathrm{NE_{o,t-1,z}}$$

[100] See Goldsmith-Pinkham, Sorkin and Swift (2020).

where each NE is a count of employed persons (see Section 3.2). The above is an accounting identity, as

$$\mathrm{X_t + X_{o,t} - X_t + X_{o,t,z} - X_{o,t} = X_{o,t,z}}$$

and

$$\sum_{\mathrm{o}} \mathrm{(X_{o,t,z}) \cdot NE_{o,t,z}/NE_{t,z} = X_{t,z}}$$

Is the Classical Accounting Identity (see Section 3.2).To arrive at an instrument, Bartik *removes* the "Shift Component" - because this is the component most sensitive to endogeneity - and also lags the Shares $\mathrm{NE_{o,t,z}/NE_{t,z}}$ to a base period t:[101]

$$\begin{aligned}\tilde{\mathrm{X}}_{\mathrm{t,z}} &= \sum_{\mathrm{o}} \Big(\overbrace{\mathrm{X_t}}^{\substack{\text{“National}\\ \text{Component”}}} + \overbrace{\mathrm{X_{o,t} - X_t}}^{\substack{\text{“Share}\\ \text{Component”}}} \Big) \cdot \mathrm{NE}_{\mathrm{o},t,\mathrm{z}}/\mathrm{NE}_{t,\mathrm{z}} \\ &= \sum_{\mathrm{o}} (\mathrm{X_{o,t}}) \cdot \mathrm{NE}_{\mathrm{o},t,\mathrm{z}}/\mathrm{NE}_{t,\mathrm{z}}\end{aligned}$$

I posit that there is some redundancy in Bartik's terminology. Although Bartik refers to,

$$\mathrm{(X_{o,t} - X_t) \cdot NE_{o,t,z}/NE_{t,z}}$$

as the "Share Component," the Share itself is surely just,

$$\mathrm{NE_{o,t,z}/NE_{t,z}}$$

[101] The formula actually used by Bartik (1991) may be different. It appears that he lags Localities' Shares of Industries rather than Industries' Shares of Localities (see Section 3.2), but does not provide a rationale for this choice. Localities' Shares of Industries can be part of the same accounting identity in this case, only because the variable X itself contains NE terms, which can be exchanged with those of the natural Shares. Rather, the interpretation I present here is applicable to any variable X, and is truer to how most researchers have understood Bartik instruments subsequently.

That which Bartik calls the "Shift Component,"

$$(\mathrm{X}_{\mathrm{o,t,z}} - \mathrm{X}_{\mathrm{o,t}}) \cdot \mathrm{NE}_{\mathrm{o,t,z}}/\mathrm{NE}_{\mathrm{t,z}}$$

also contains the Share. I call rather $\mathrm{X}_{\mathrm{o,t,z}}$ the Shift, and $\mathrm{X}_{\mathrm{o,t}}$ the Delocalized Shift. In either case, the steps for converting the accounting identity into the instrumental variable are equivalent, only said in a different way. That is, *removal* of what Bartik calls the "Shift Component" is equivalent to *delocalization* of what I call the Shift. In either case, the accounting identity can be written as,

$$\mathrm{X}_{\mathrm{t,z}} = \sum_{\mathrm{o}} \mathrm{X}_{\mathrm{o,t,z}} \cdot \mathrm{NE}_{\mathrm{o,t,z}}/\mathrm{NE}_{\mathrm{t,z}}$$

and the instrument is,

$$\tilde{\mathrm{X}}_{\mathrm{t,z}} = \sum_{\mathrm{o}} \mathrm{X}_{\mathrm{o,t}} \cdot \mathrm{NE}_{\mathrm{o},\mathit{t},\mathrm{z}}/\mathrm{NE}_{\mathit{t},\mathrm{z}}$$

The main step in converting $\mathrm{X}_{\mathrm{t,z}}$ to $\tilde{\mathrm{X}}_{\mathrm{t,z}}$ is the Shift Delocalization, $\mathrm{X}_{\mathrm{o,t,z}} \longrightarrow \mathrm{X}_{\mathrm{o,t}}$, and the secondary step is the Share Lag, $\mathrm{NE}_{\mathrm{o,t,z}}/\mathrm{NE}_{\mathrm{t,z}} \longrightarrow \mathrm{NE}_{\mathrm{o},\mathit{t},\mathrm{z}}/\mathrm{NE}_{\mathit{t},\mathrm{z}}$.

Bound and Holzer (2000), also closely related to Bartik (1991) and Blanchard and Katz (1992), study in particular how the impacts of local job growth fall differently across racial and demographic groups. Their Shift-Share instrument is also the same,[102] only with a very slight context alteration. Where Bartik and Blanchard and Katz view job growth in terms of the total number of people working, Bound and Holzer (2000) view it rather in terms of the total number of hours worked. In other words, where Bartik would view each individual as either 1 (a worker) or 0 (not a worker), Bound and Holzer would view a part-time worker as effectively 1/2.

Although only slightly different from Bartik (1991), the case of Bound and Holzer (2000) provides a simple illustration of how the Shift-Share instrument may adapt

[102] That is, they are the same under my interpretation in which Shares are Industries' Shares of Localities. Bound and Holzer (2000) appear indeed to view Shares as Industries' Shares of Localities, although they do not clarify this explicitly.

while retaining the same structure and properties. Because the endogenous explanatory variable of interest $\mathrm{X_{t,z}}$ for Bound and Holzer is the (change over time in) total hours worked, the Shift $\mathrm{X_{o,t,z}}$ in their accounting identity is also the (change over time in) total hours worked - by industry o.

Classical Accounting Identity:

$$\mathrm{X_{t,z}} = \sum_{\mathrm{o}} \overbrace{\mathrm{X_{o,t,z}}}^{\text{Shift}} \cdot \overbrace{\mathrm{NE_{o,t,z}} \Big/ \mathrm{NE_{t,z}}}^{\text{Share}}$$

Shift-Share Instrument:

$$\tilde{\mathrm{X}}_{\mathrm{t,z}} = \sum_{\mathrm{o}} \overbrace{\mathrm{X_{o,t}}}^{\text{Shift}} \cdot \overbrace{\mathrm{NE}_{\mathrm{o},t,\mathrm{z}} \Big/ \mathrm{NE}_{t,\mathrm{z}}}^{\text{Share}}$$

Both the accounting identity and instrument are the same as in the original case. Only the content of the explanatory variable $\mathrm{X_{t,z}}$ has changed - from a measure of (change over time in) total number of workers, to a measure of (change over time in) total number of hours worked.

The classical accounting identity and instrument are applicable for any endogenous explanatory variable $\mathrm{X_{t,z}}$ that can be decomposed over an industrial, occupation, or similar categorization o.[103] Rather than job growth, Diamond (2016) for example studies wages as explanatory variables $\mathrm{X_{t,z}}$. Therefore, the Delocalized Shift $\mathrm{X_{o,t}}$ for Diamond refers to industrial average wages - whereas for Bound and Holzer (2000) and others it referred to industrial average job growth.

[103]The essential feature of such a categorization is that the Share vector it defines should be less sensitive to endogeneity than is the explanatory variable itself.

3.4 Essential and Adjustable Features

It is worth emphasizing that Shift-Share instruments, even in the classical setup, contain multiple features that may be adjusted without straying from the essential properties and purpose. This section discusses some of these features, partly in order to clarify which aspects of classical structure are most essential.

3.4.1 Lagging of Shares

One of the core properties of Shift-Share instruments is that the Share vector retains some information that is distinctive to its cross sectional unit or Locality, z. By contrast, the Shift vector is delocalized over the cross section for the sake of exogeneity, severing its connection to z. But the precise manner in which Share component retains its connection to z may be adjusted. Typically it is lagged to a historical base period; but it may be lagged in other ways. Moreover, because the Share vector is meant to retain that which is most exogenous about the regressor per z, it is conceivable that the Shares needn't be lagged at all.

Most commonly, and including in Bartik (1991), the Share vector is lagged to a historical base period t. That is, the Share component of the regressor $\mathrm{X}_{\mathrm{t,z}}$ is replaced as $(\mathrm{z}, \mathrm{t}) \longrightarrow (\mathrm{z}, t)$. I call this a Frozen Lag: All t are lagged to a fixed t, regardless of the distance in time between t and t. In Bartik (1991), t is 1970, while $\mathrm{t} \in [1971, 1986]$ for the panel itself. Critically, this lagged Share vector (of industrial shares) is assumed to be causally prior to (exogenous with respect to) future job growth. The main argument for this exogeneity is that industrial shares are deep characteristics of localities (MSAs) z, arising from forces outside the model, such as geography, technology, and historical accident. The lag provides an auxiliary argument, applicable but not conclusive (on its own) in any context: Anything in the past is likely to be causally to prior to anything in the present or future.

I argue that, although less common in Shift-Share instruments than Frozen Lags are, Updating Lags are usually better. In contrast to Frozen Lags, which lag all t to a common base period t, lagging $(\mathrm{z}, \mathrm{t}) \longrightarrow (\mathrm{z}, t)$, Updating Lags rather maintain a common distance in time between each t and its lag, $(\mathrm{z}, \mathrm{t}) \longrightarrow (\mathrm{z}, \mathrm{t}-1)$. As discussed in the previous paragraph, the purpose of either kind of lag is to help insure that the Shares are causally prior to the dependent variable in the present (t), although this is not the main argument for the Shares' exogeneity. Either kind of lag accomplishes this purpose equally, as either is prior in time to the present (t). However, Frozen Lags have the downside that because the base period t is a different distance in time from each t in the panel, it is likely to be far more relevant for the earlier time periods - those closer to t - than it is for the later time periods. The magnitude of this problem is likely to depend on the span of time under study.

In addition to Frozen or Updating Lags, Shares may be averaged over all lags and leads, or not lagged at all. Averaging over all lags and leads may be sensible particularly if the number of time periods is large, as in Nunn and Qian (2014). With a large number of time periods, that is, the role of any one time period's values in determining the averages for z is small. Hence these averages can be assumed to reflect intrinsic characteristics of z, causal first movers rather than endogenous reactions. However, perhaps most thematically important to consider is the option of not lagging the Shares at all. As discussed earlier in this section, the lag (of whichever kind) is in any case not the primary argument for the Shares' exogeneity. Rather, the primary argument is that Shares are deep characteristics of localities z, arising from forces outside the model - such as geography, technology and historical accident. This primary argument may be applicable in the present (t) as well as in any lag.

3.4.2 Differencing and Fixed Effects

Although Shift-Share instruments provide a powerful framework for bypassing reverse causality, it should be noted that they do not offer anything novel with respect to omitted variable bias. Rather, Shift-Share instruments rely on standard methods for unobservables: differencing, fixed effects, and other controls. Differencing and fixed effects are methods meant to account for the sum total of all unobserved effects. However, neither of these is perfect, and the potential for unobserved information to play a confounding role is ultimately unavoidable in any setting.

The stereotypical Shift-Share instrument is accompanied by a differencing method. This can be understood as follows. Suppose the model is of a form,

$$\mathrm{Y}_{\mathrm{t,z}} = \beta \cdot \mathrm{X}_{\mathrm{t,z}} + \delta_{\mathrm{t}} + \delta_{\mathrm{z}} + \delta_{\mathrm{t,z}} \tag{49}$$

where $\mathrm{X}_{\mathrm{t,z}}$ is the endogenous regressor of interest, and δ are unobserved factors. As we know, the endogenous regressor $\mathrm{X}_{\mathrm{t,z}}$ can be written as a product of Shift and Share components (accounting identity),

$$\mathrm{X}_{\mathrm{t,z}} = \sum_{\mathrm{o}} \overbrace{\mathrm{X}_{\mathrm{o,t,z}}}^{\text{Shift}} \cdot \overbrace{\mathrm{N}_{\mathrm{o,t,z}} \big/ \mathrm{N}_{\mathrm{t,z}}}^{\text{Share}} \tag{50}$$

The Shift-Share instrument for $\mathrm{X}_{\mathrm{t,z}}$,

$$\tilde{\mathrm{X}}_{\mathrm{t,z}} = \sum_{\mathrm{o}} \overbrace{\mathrm{X}_{\mathrm{o,t}}}^{\text{Shift}} \cdot \overbrace{\mathrm{N}_{\mathrm{o},t,\mathrm{z}} \big/ \mathrm{N}_{t,\mathrm{z}}}^{\text{Share}} \tag{51}$$

corrects for reverse causality, but not necessarily for bias arising from the unobserved factors δ.

Stereotypically, a time difference of (49) is taken to eliminate the time-invariant unobserved component, δ_{z}. The differenced equation is,

$$(\mathrm{Y}_{\mathrm{t,z}} - \mathrm{Y}_{t,\mathrm{z}}) = \beta \cdot (\mathrm{X}_{\mathrm{t,z}} - \mathrm{X}_{t,\mathrm{z}}) + (\delta_{\mathrm{t}} - \delta_{t}) + (\delta_{\mathrm{t,z}} - \delta_{t,\mathrm{z}}) \tag{52}$$

with

$$\tilde{\mathrm{X}}_{\mathrm{t,z}} - \tilde{\mathrm{X}}_{t,\mathrm{z}} = \sum_{\mathrm{o}} \overbrace{(\mathrm{X}_{\mathrm{o,t}} - \mathrm{X}_{\mathrm{o},t})}^{\text{Shift}} \cdot \overbrace{\mathrm{N}_{\mathrm{o},t,\mathrm{z}} \big/ \mathrm{N}_{t,\mathrm{z}}}^{\text{Share}}$$

straightforwardly serving as the instrument for $(\mathrm{X}_{\mathrm{t,z}} - \mathrm{X}_{t,\mathrm{z}})$.[104] Estimation of (52) can recover β without fear of confounding from δ_{z}, because the time difference has eliminated δ_{z}. Or, rather than taking the time difference, an alternative method to accomplish the same purpose is to impose fixed effects $\hat{\delta}_{\mathrm{z}}$ to absorb δ_{z} in the estimation of (49). In either case, time period fixed effects $\hat{\delta}_{\mathrm{t}}$ are imposed as well to absorb δ_{t}.

A weakness of either differencing or fixed effects $\hat{\delta}_{\mathrm{z}}$ is that, although both nullify time-invariant unobservables δ_{z} completely, neither does anything whatsoever to nullify time-varying unobservables $\delta_{\mathrm{t,z}}$. There is typically no reason to assume that all potentially confounding unobservables would be time-invariant, and hence time-varying unobservables are the most important blind spot of many Shift-Share instruments. Alternatively, time-varying regional effects $\hat{\delta}_{\mathrm{t,Z}}$ may be imposed to absorb some time-varying unobservables, as well as some time-invariant unobservables. (Regions Z are closely related clusters of Localities z.)[105] The optimal balance of such effects should absorb the most important potentially confounding unobserved factors of both types, that is, time-varying and time-invariant.

3.5 Novel Variants

Autor, Dorn and Hanson (2013) study the effects of import competition on labor market outcomes for US workers, that is, unemployment, labor force participation, and wages. The endogenous explanatory variable of interest is a change in labor market exposure to Chinese imports. Total China-to-US import volume is I_{t}. The basic idea is that these imports I_{t} are in competition with the output of US workers:

[104] Often $(\mathrm{X}_{\mathrm{t,z}} - \mathrm{X}_{t,\mathrm{z}})$ is written rather as the original $\mathrm{X}_{\mathrm{t,z}}$, and $(\mathrm{X}_{\mathrm{o,t}} - \mathrm{X}_{\mathrm{o},t})$ as the original $\mathrm{X}_{\mathrm{o,t}}$.

[105] Time-varying effects at the z level $\hat{\delta}_{\mathrm{t,z}}$ would absorb the entire panel, leaving nothing to be explained by any regressor(s) of interest.

The competitive impact of $\mathrm{I_t}$ on US workers is spread out equally, so that the impact per individual US worker is $\mathrm{I_t/NE_t}$, where $\mathrm{NE_t}$ is the total number of workers in the US. The change in exposure can be written as

$$\mathrm{X_{t,z} = (I_t - I_{t-1})/NE_{t-1}} \tag{53}$$

for any US Commuting Zone z in which there are any workers. This $\mathrm{X_{t,z}}$ is constant over z, assuming there are any workers in z, because the impact is assumed to be spread equally over workers nation-wide.

Instead of (53), which is an oversimplification, the authors suppose rather that the competitive impact of Chinese imports *in industry o* is spread equally nation-wide over all US workers *in industry o*.

$$\mathrm{X_{o,t,z} = (I_{o,t} - I_{o,t-1})/NE_{o,t-1}} \tag{54}$$

Alhough specific by industry o, $\mathrm{X_{o,t,z}}$ is still constant over z, for the same reason that equation (53)'s $\mathrm{X_{t,z}}$ would be constant over z. However, the unconditional average impact per worker in z now depends on the industrial makeup of z, because each industry-specific impact $\mathrm{X_{o,t,z}}$ applies only to the workers in industry o:

$$\mathrm{X_{t,z}} = \sum_{\mathrm{o}} \overbrace{\mathrm{X_{o,t,z}}}^{\text{Shift}} \cdot \overbrace{\mathrm{NE_{o,t,z}}\Big/\mathrm{NE_{t,z}}}^{\text{Share}} \tag{55}$$

$$\mathrm{X_{o,t,z} = (I_{o,t} - I_{o,t-1})/NE_{o,t-1}}$$

Equation (55) happens to have (apparently) the same form as the Classical Accounting Identity (see Section 3.2). And, as in the classical setting, Autor, Dorn and Hanson (2013) use lagged shares in their instrument, $\mathrm{NE_{o,t-1,z}}\Big/\mathrm{NE_{t-1,z}}$ in place of $\mathrm{NE_{o,t,z}}\Big/\mathrm{NE_{t,z}}$. Yet, the Shift $\mathrm{X_{o,t,z}}$ is unlike that in the classical setting, because it is already constant over z. Typically, a Shift Delocalization involves replacing local

Industry averages $X_{o,t,z}$ with national Industry averages $X_{o,t}$. However, as discussed in the previous paragraph, it is already the case that $X_{o,t,z} = X_{o,t}$, due to the way in which the exposure variable itself is defined. But Autor, Dorn and Hanson (2013) go further than this, delocalizing their Shift on a higher level. They replace *nation*-wide industry averages $X_{o,t}$ with *world*-wide averages $X'_{o,t}$.

$$\tilde{X}_{t,z} = \sum_{o} \overbrace{X'_{o,t}}^{\text{Shift}} \cdot \overbrace{NE_{o,t-1,z} \Big/ NE_{t-1,z}}^{\text{Share}} \tag{56}$$

$$X'_{o,t} = (I'_{o,t} - I'_{o,t-1})/NE_{o,t-1}$$

I'_o is export volume from China industry o *world*-wide, to other high income countries besides the US.

Nunn and Qian (2014) study the effect of food aid on armed conflict in countries receiving the aid. Although food aid might be expected to cool tensions between opposing factions, often rather the opposite is observed, as armed theft of the aid ignites further conflict. There is also a strong channel for reverse causality, however: Countries experiencing conflict may be more likely to receive aid.

The endogenous variable of interest $X_{t,z}$ for Nunn and Qian is the the amount of food aid (wheat) that country z receives from the US in year t. An accounting identity for this variable is,

$$X_{t,z} = \overbrace{I_{t,z}}^{\text{Share}} \cdot \overbrace{X_{t-1}}^{\text{Push}} \tag{57}$$

where $I_{t,z}$ is a binary variable equal to one if country z is selected to receive wheat aid in year t, and X_{t-1} is the quantity of wheat produced in the US in the previous year (scaled by the number of countries to receive the aid). This quantity is lagged because it takes a one-year cycle for the wheat to transition from production to distribution.

Their instrument is,

$$\tilde{\mathrm{X}}_{\mathrm{t,z}} = \overbrace{T^{-1}\sum_{t} I_{t,\mathrm{z}}}^{\text{Share}} \cdot \overbrace{\mathrm{X}_{\mathrm{t-1}}}^{\text{Push}} \tag{58}$$

accompanied by region specific time trends and country fixed effects. The summation index t covers all years in the sample, and $T = 36$ is the number of years t. Observations are 125 non-OECD countries, over the 36 years. The outcome variable is a binary indicator of whether the country is in a state of conflict, defined as experiencing more than 25 battle deaths in the year.

Nunn and Qian's instrument is similar to a classical Shift-Share instrument in the Share, except that they use an unusual kind of share, and lag it in an unusual way. Rather than an Updating or Frozen Lag (see Section 3.4.1), they take the average over all lags and leads. This is appropriate both because the number of time periods is large, and because of the unusual binary nature of the Share. Because the number of time periods is large, the effect of any one time period's value in determining the average is small. Therefore the time averaged Share can be viewed as a deep characteristic of the country z (conditional on country fixed effects), rather than an endogenous reaction to present period conditions.

Because $\mathrm{X}_{\mathrm{t-1}}$ is already constant over countries z, it cannot be delocalized as a Shift would be. I therefore refer to it as something else, a Push. I define a Push as a component of the accounting identity that is not altered between the accounting identity and the instrument. Many Shift-Share instruments contain such components. But would I not refer to Nunn and Qian's instrument as a Shift-Share instrument per se, because it does not contain a Shift that is delocalized for the purpose of enhancing exogeneity.

Card (2001) studies the effects of immigrant inflows on occupation specific wages, employment and unemployment rates of natives. A larger pool of workers competing for the same set of jobs j may depress wages by diminishing laborers' bargaining power against employers. However, this effect is difficult to isolate because, for example, cities with higher wages may attract more immigrant workers.

For each occupation group j, Card's endogenous explanatory variable of interest $\mathrm{X_{j,t,z}}$ is the inflow of immigrant workers of group j, $\mathrm{NI_{j,t,z}}$, to each Metropolitan Statistical Area (MSA) z in the US. This inflow can be expressed as the sum of inflows from each of various countries of origin, o.

$$\mathrm{X_{j,t,z}} = \mathrm{NI_{j,t,z}} = \sum_{\mathrm{o}} \mathrm{NI_{j,o,t,z}}$$

This can be multiplied by $\mathrm{NI_{o,t,z}/NI_{o,t,z}}$ and $\mathrm{NI_{o,t}/NI_{o,t}}$ to yield the following accounting identity:

$$\mathrm{X_{t,z}} = \sum_{\mathrm{o}} \overbrace{\mathrm{NI_{j,o,t,z}/NI_{o,t,z}}}^{\text{Shift}} \cdot \overbrace{\mathrm{NI_{o,t,z}}\Big/\mathrm{NI_{o,t}}}^{\text{Share}} \cdot \overbrace{\mathrm{NI_{o,t}}}^{\text{Push}} \tag{59}$$

$\mathrm{NI_{o,t}}$ is the total inflow to the US of immigrant workers from country of origin o. $\mathrm{NI_{o,t,z}/NI_{o,t}}$ are the fraction of immigrant workers from country of origin o who accrue to each MSA z. Abstractly, $\mathrm{NI_{o,t,z}/NI_{o,t}}$ are Localities' Shares of Industries (see Section 3.2), with countries of origin o as abstract Industries. $\mathrm{NI_{j,o,t,z}/NI_{o,t,z}}$ are the fraction of immigrant workers from country of origin o to MSA z who are workers in occupation group j.

To arrive at an instrument, Card delocalizes the Shifts $\mathrm{NI_{j,o,t,z}/NI_{o,t,z}}$ over z, and lags the Shares $\mathrm{NI_{o,t,z}/NI_{o,t}}$ to a historical base period t.[106]

$$\tilde{\mathrm{X}}_{\mathrm{t,z}} = \sum_{\mathrm{o}} \underbrace{\overbrace{\mathrm{NI_{j,o,t}/NI_{o,t}}}^{\text{Shift}}}_{\tau_{\mathrm{gj}}} \cdot \underbrace{\overbrace{\mathrm{NI}_{\mathrm{o},t,\mathrm{z}}\Big/\mathrm{NI}_{\mathrm{o},t}}^{\text{Share}}}_{\lambda_{\mathrm{gc}}} \cdot \underbrace{\overbrace{\mathrm{NI_{o,t}}}^{\text{Push}}}_{\mathrm{M_g}}$$

Similarly as in a classical Shift-Share instrument, the Shares represent that which is most exogenous about $\mathrm{X_{t,z}}$ per Locality z. Newly arriving immigrants tend to move

[106] τ_{gj}, λ_{gc}, $\mathrm{M_g}$ are the notation actually used by Card (2001) for these objects.

to enclaves established by earlier immigrants from the same source country, not for endogenous labor demand reasons, but rather due to cultural and family ties. The lag provides an additional guarantee of this exogeneity.

Also like in a classical Shift-Share instrument, the local Shifts would be highly endogenous, and therefore are delocalized over the cross section. These local Shifts $\mathrm{NI}_{\mathrm{j,o,t,z}}/\mathrm{NI}_{\mathrm{o,t,z}}$ give the occupational makeup of immigrant workers entering each city. Occupational makeup is driven heavily by the local economy's labor demand, which is endogenous in this setting. The Delocalized Shifts $\mathrm{NI}_{\mathrm{j,o,t}}/\mathrm{NI}_{\mathrm{o,t}}$ replace the occupational makeup entering each city with the average occupational makeup of immigrants entering the US as a whole, by country of origin o.

Boustan et al. (2013) study the effects of income inequality on public taxation and expenditure. On the one hand, a highly unequal local population may have little in common with one another, and therefore little that they can agree on in the way of public programs. On the other hand, median voter theory suggests that a more unequal distribution will yield a more covetous median voter, and therefore a larger appetite for public programs. The authors show that the latter effect dominates. But to do so, they must get around the reverse causal effect of public programs on inequality.

Boustan et al. (2013) build an instrument for the Gini coefficient of income, a traditional measure of inequality. With income data for individuals i, the Gini coefficient in any given Locality (school district) z in time period t would be,

$$\mathrm{X}_{\mathrm{t,z}} = \frac{\sum_{i\in \mathrm{t,z}} \sum_{\mathrm{i}\in \mathrm{t,z}} \left| \mathrm{Y}_{\mathrm{i,t,z}} - \mathrm{Y}_{i,\mathrm{t,z}} \right|}{2\mathrm{N}_{\mathrm{t,z}} \sum_{i\in \mathrm{t,z}} \mathrm{Y}_{i,\mathrm{t,z}}} \tag{60}$$

where $\mathrm{Y}_{\mathrm{i,t,z}}$ or $\mathrm{Y}_{i,\mathrm{t,z}}$ is the income of any given individual i or i. For data reasons, the authors use rather a discretized Gini, that is, an approximation of the Gini using discrete bins h. Concretely, whatever bin h that $\mathrm{Y}_{\mathrm{i,t,z}}$ falls into, $\mathrm{Y}_{\mathrm{i,t,z}}$ is discretized rather to be equal to $\mathrm{Y_{h,t}}$, a midpoint value for the bin.[107]

$$\mathrm{X_{t,z}} = \frac{\sum_{h}\sum_{\mathrm{h}} \overbrace{\mathrm{P_{h,t,z}P}_{h,\mathrm{t,z}}}^{\text{Share}} \cdot \overbrace{\left|\mathrm{Y_{h,t}} - \mathrm{Y}_{h,\mathrm{t}}\right|}^{\text{Push}}}{2\sum_{h} \underbrace{\mathrm{P}_{h,\mathrm{t,z}}}_{\text{Share}} \cdot \underbrace{\mathrm{Y}_{h,\mathrm{t}}}_{\text{Push}}}, \quad \mathrm{P_{h,t,z}} = \mathrm{N_{h,t,z}/N_{t,z}} \tag{61}$$

$\mathrm{P_{h,t,z}} = \mathrm{N_{h,t,z}/N_{t,z}}$, which I call Bin Shares, are the fractions of the local z population who fall into each bin h. The cutoffs of the bins h are shared in common over the nation, so the Bin Shares vary by Locality z. For example, more unequal Localities z will tend to have higher $\mathrm{P_{h,t,z}}$ values for the highest and lowest bins h. To arrive at an instrument, the authors lag the Bin Shares to a historical base period t.

$$\tilde{\mathrm{X}}_{\mathrm{t,z}} = \frac{\sum_{h}\sum_{\mathrm{h}} \overbrace{\mathrm{P}_{\mathrm{h},t,\mathrm{z}}\mathrm{P}_{h,t,\mathrm{z}}}^{\text{Share}} \cdot \overbrace{\left|\mathrm{Y_{h,t}} - \mathrm{Y}_{h,\mathrm{t}}\right|}^{\text{Push}}}{2\sum_{h} \underbrace{\mathrm{P}_{h,t,\mathrm{z}}}_{\text{Share}} \cdot \underbrace{\mathrm{Y}_{h,\mathrm{t}}}_{\text{Push}}} \tag{62}$$

They also use a time difference, plus linear trend effects and other controls, to absorb unobservables. Because there is no Shift vector that is delocalized over the cross section in moving from the accounting identity to the instrument, I would not call it a Shift-Share instrument per se - although it may nonetheless be a valid instrument. Using this example specifically, I discuss what I view as the potential importance of this distinction in depth in Section 3.7.

3.6 New Creations

The abstracted Shift-Share instrument is a flexible framework for constructing potential instrumental variables. I posit that the core idea of the Shift-Share approach

[107] For example, if the cutoffs for bin a bin h are \$30,000 on the low end and \$34,000 on the high end, an individual i's income $\mathrm{Y_{i,t,z}}$ = \$30,500 may be discretized as $\mathrm{Y}_{h,\mathrm{t}}$ = \$32,000.

is to decompose the endogenous explanatory variable as an accounting identity that contains a more endogenous factor (the Shift vector) and a more exogenous factor (the Share vector); replace the Shift vector with an analogue that is completely exogenous, because it is delocalized over the cross section; and lag the Share vector to enhance its already solid claim to exogeneity. The Shift Delocalization and Share Lag - in conjunction with differencing methods and controls for unobservables - provide Shift-Share instruments' claim to exogeneity. Conditional on these, starting from an accounting identity provides a guarantee that the instrument should be relevant.

In Section 3.2, I presented a general formula for classical Shift-Share instruments, applicable to any endogenous explanatory variable $\mathrm{X_{t,z}}$ which is a mean over individuals who can be rearranged into subgroups to form Shares. In this section, I derive analogous formulas for other distribution summaries, that is, besides the mean. As with the formula for the mean, basing the instrument on an accounting identity is a guide to relevance. That is, it provides the most natural approximation of the endogenous explanatory variable as a function of the exogenous factors.[108]

3.6.1 Variance and Skew

In Section 3.2, I considered any endogenous explanatory variable $\mathrm{X_{t,z}}$ that is a mean over individuals i. Following from the definitions of means and from the fact that,

$$\sum_{\mathrm{i \in t,z}} \mathrm{X_i} = \sum_{\mathrm{o}} \sum_{\mathrm{i \in o,t,z}} \mathrm{X_i} \tag{63}$$

I showed that $\mathrm{X_{t,z}}$ can be expressed as the Classical Accounting Identity,

[108] A first stage consisting of all the lagged Shares as separate instruments, for example, may yield a higher overall fit - but this would most likely be a case of overfitting.

$$X_{t,z} = \sum_{o} \overbrace{X_{o,t,z}}^{\text{Shift}} \cdot \overbrace{N_{o,t,z}/N_{t,z}}^{\text{Share}}$$

I now consider any endogenous explanatory variable $V_{t,z}$ that is a variance over individuals i.

$$V_{t,z} = N_{t,z}^{-1} \sum_{i \in t,z} (X_i - X_{t,z})^2 \tag{64}$$

For readability, I suppress t subscripts throughout the remainder of Section 3.6.

$$V_z = N_z^{-1} \sum_{i \in z} (X_i - X_z)^2 \;\Leftrightarrow\; \sum_{i \in z} (X_i - X_z)^2 = N_z \cdot V_z \tag{65}$$

In addition to (63), it is helpful to apply the identity,

$$(X_i - X_z)^2 = (X_i - X_{o,z})^2 + ((X_i - X_z)^2 - (X_i - X_{o,z})^2) \tag{66}$$

(66) beneficially decomposes the moment into within-category and across-category components.

$$\begin{aligned} V_z &= N_z^{-1} \sum_{o} \sum_{i \in o,z} (X_i - X_z)^2 \\ &= N_z^{-1} \sum_{o} \Big\{ \sum_{i \in o,z} \underbrace{(X_i - X_{o,z})^2}_{\text{within–o}} + \sum_{i \in o,z} \underbrace{((X_i - X_z)^2 - (X_i - X_{o,z})^2)}_{\text{across–o}} \Big\} \end{aligned} \tag{67}$$

The within-category component is straightforward, as by definition,

$$V_{o,z} = N_{o,z}^{-1} \sum_{i \in o,z} (X_i - X_z)^2 \;\Leftrightarrow\; \sum_{i \in o,z} (X_i - X_{o,z})^2 = N_{o,z} \cdot V_{o,z} \tag{68}$$

The across-category component simplifies after expansion and collection of terms:

$$\begin{aligned}
&\sum_{i\in o,z}((X_i - X_z)^2 - (X_i - X_{o,z})^2) \\
&= \sum_{i\in o,z}(X_i^2 - 2X_iX_z + X_z^2 - X_i^2 + 2X_iX_{o,z} - X_{o,z}^2) \\
&= 2(X_{o,z} - X_z)\sum_{i\in o,z}(X_i) - (X_{o,z}^2 - X_z^2)\sum_{i\in o,z}(1) \\
&= 2(X_{o,z} - X_z)\cdot N_{o,z}\cdot X_{o,z} - (X_{o,z}^2 - X_z^2)N_{o,z} \\
&= 2(X_{o,z}^2 - X_{o,z}X_z)\cdot N_{o,z} - (X_{o,z}^2 - X_z^2)\cdot N_{o,z} \\
&= (X_{o,z}^2 - 2X_{o,z}X_z + X_z^2)\cdot N_{o,z} \\
&= N_{o,z}\cdot(X_{o,z} - X_z)^2
\end{aligned} \tag{69}$$

Combining the above,

$$V_z = \sum_o \overbrace{\left\{ \underbrace{V_{o,z}}_{\substack{\text{within–o}\\ \text{variance}}} + \underbrace{(X_{o,z} - X_z)^2}_{\substack{\text{across–o}\\ \text{variance}}} \right\}}^{\text{Shift}} \cdot \overbrace{N_{o,z}/N_z}^{\text{Share}} \tag{70}$$

Equation (70) is the analogue of the Classical Accounting Identity for variances rather than means. With time period subscripts, this is:

$$V_{t,z} = \sum_o \overbrace{\left\{ \underbrace{V_{o,t,z}}_{\substack{\text{within–o}\\ \text{variance}}} + \underbrace{(X_{o,t,z} - X_{t,z})^2}_{\substack{\text{across–o}\\ \text{variance}}} \right\}}^{\text{Shift}} \cdot \overbrace{N_{o,t,z}/N_{t,z}}^{\text{Share}} \tag{71}$$

To arrive at the Shift-Share instrument, delocalize the Shift vector over z, and lag the Share vector:

$$\tilde{V}_{t,z} = \sum_{o} \overbrace{\left\{ \underbrace{V_{o,t}}_{\substack{\text{within–o}\\ \text{variance}}} + \underbrace{(X_{o,t} - \tilde{X}_{t,z})^2}_{\substack{\text{across–o}\\ \text{variance}}} \right\}}^{\text{Shift}} \cdot \overbrace{N_{o,t,z}/N_{t,z}}^{\text{Share}} \tag{72}$$

where $\tilde{X}_{t,z}$ is the Shift-Share instrument for the mean.

The above equations illustrate the importance of deriving the instrument from an accounting identity, rather than from intuition alone. Intuitively, one might construct an instrument for the variance using only the within category variance,

$$\tilde{V}_{t,z} = \sum_{o} \overbrace{V_{o,t}}^{\text{Shift}} \cdot \overbrace{N_{o,t,z}/N_{t,z}}^{\text{Share}} \tag{73}$$

because this may seem most analogous to the Shift-Share instrument for the mean,

$$\tilde{X}_{t,z} = \sum_{o} \overbrace{X_{o,t}}^{\text{Shift}} \cdot \overbrace{N_{o,t,z}/N_{t,z}}^{\text{Share}} \tag{74}$$

Similarly as (74) is a Share-weighted average of the category (Industry) o's national averages, (73) would be a Share-weighted average of the Industries' national variances. Indeed, one might conceive of constructing an instrument for any kind of moment in this way. But the accounting identity (71) demonstrates that there is additional term, the across-category variance, that should be included as well in order to arrive at the most natural prediction of $V_{t,z}$ as a function of the national industry distributions.

To derive the instrument for the skew,

$$W_z = N_z^{-1} \sum_{i \in z} (X_i - X_z)^3$$

it is again helpful to decompose into within and across category components,

$$(X_i - X_z)^3 = (X_i - X_{o,z})^3 + ((X_i - X_z)^3 - (X_i - X_{o,z})^3)$$

The across category component now simplifies as,

$$
\begin{aligned}
&\sum_{\mathrm{i}\in\mathrm{o,z}}((\mathrm{X_i}-\mathrm{X_z})^3-(\mathrm{X_i}-\mathrm{X_{o,z}})^3)\\
&=\sum_{\mathrm{i}\in\mathrm{o,z}}(\mathrm{X_i^3}-3\mathrm{X_i^2X_z}+3\mathrm{X_iX_z^2}-\mathrm{X_z^3}-\mathrm{X_i^3}+3\mathrm{X_i^2X_{o,z}}-3\mathrm{X_iX_{o,z}^2}+\mathrm{X_{o,z}^3})\\
&=3(\mathrm{X_{o,z}}-\mathrm{X_z})\sum_{\mathrm{i}\in\mathrm{o,z}}(\mathrm{X_i^2})-3(\mathrm{X_{o,z}^2}-\mathrm{X_z^2})\sum_{\mathrm{i}\in\mathrm{o,z}}(\mathrm{X_i})+(\mathrm{X_{o,z}^3}-\mathrm{X_z^3})\sum_{\mathrm{i}\in\mathrm{o,z}}(1)\\
&=3(\mathrm{X_{o,z}}-\mathrm{X_z})\cdot\mathrm{N_{o,z}}\cdot(\mathrm{V_{o,z}}+\mathrm{X_{o,z}^2})-3(\mathrm{X_{o,z}^2}-\mathrm{X_z^2})\cdot\mathrm{N_{o,z}}\cdot\mathrm{X_{o,z}}+(\mathrm{X_{o,z}^3}-\mathrm{X_z^3})\cdot\mathrm{N_{o,z}}\\
&=(\mathrm{X_{o,z}^3}-3\mathrm{X_{o,z}^2X_z}+\mathrm{X_{o,z}X_z^2}-\mathrm{X_z^3})\cdot\mathrm{N_{o,z}}+3(\mathrm{X_{o,z}}-\mathrm{X_z})\cdot\mathrm{N_{o,z}}\cdot\mathrm{V_{o,z}}\\
&=\mathrm{N_{o,z}}\cdot((\mathrm{X_{o,z}}-\mathrm{X_z})^3+3(\mathrm{X_{o,z}}-\mathrm{X_z})\cdot\mathrm{V_{o,z}})
\end{aligned}
$$

That is, the across category component itself now contains two sub-components.

$$
\mathrm{W_z}=\sum_{\mathrm{o}}\overbrace{\left\{\underbrace{\mathrm{W_{o,z}}}_{\substack{\text{within–o}\\\text{skew}}}+\underbrace{(\mathrm{X_{o,z}}-\mathrm{X_z})^3}_{\substack{\text{across–o–mean}\\\text{skew}}}+\underbrace{3(\mathrm{X_{o,z}}-\mathrm{X_z})\mathrm{V_{o,z}}}_{\substack{\text{across–o–variance}\\\text{skew}}}\right\}}^{\text{Shift}}\cdot\overbrace{\mathrm{N_{o,z}/N_z}}^{\text{Share}}
$$

The instrument follows as,

$$
\mathrm{\tilde{W}_{t,z}}=\sum_{\mathrm{o}}\overbrace{\left\{\underbrace{\mathrm{W_{o,t}}}_{\substack{\text{within–o}\\\text{skew}}}+\underbrace{(\mathrm{X_{o,t}}-\mathrm{\tilde{X}_{t,z}})^3}_{\substack{\text{across–o–mean}\\\text{skew}}}+\underbrace{3(\mathrm{X_{o,t}}-\mathrm{\tilde{X}_{t,z}})\mathrm{V_{o,t}}}_{\substack{\text{across–o–variance}\\\text{skew}}}\right\}}^{\text{Shift}}\cdot\overbrace{\mathrm{N}_{\mathrm{o},t,\mathrm{z}}/\mathrm{N}_{t,\mathrm{z}}}^{\text{Share}}
$$

This process will solve similarly for any higher order n moment,

$$
\mathrm{N_z^{-1}}\sum_{\mathrm{i}\in\mathrm{z}}(\mathrm{X_i}-\mathrm{X_z})^{\mathrm{n}}
$$

thus allowing the researcher in principle to construct Shift-Share instruments for the entire distribution beyond the mean. I leave to the reader to derive the general n formula.

3.6.2 Mean Absolute Deviation

A mean absolute deviation around a central point X is defined as,

$$\mathrm{M_z} = \mathrm{N_z^{-1}} \sum_{\mathrm{i} \in \mathrm{z}} |\mathrm{X_i} - \mathrm{X}| \tag{75}$$

First, it is helpful to split the piecewise function into its two parts, which I call the Mean Inferior Deviation $\lfloor \mathrm{M_z}$ and Mean Superior Deviation $\lceil \mathrm{M_z}$.

$$\begin{aligned} \mathrm{M_z} &= \lfloor \mathrm{X_z} + \lceil \mathrm{X_z} \\ \lfloor \mathrm{M_z} &= \mathrm{N_z^{-1}} \sum_{\mathrm{i} \in \mathrm{z}} (\mathrm{X} - \mathrm{X_i}) \cdot I[\mathrm{X_i} \leq \mathrm{X}] \\ \lceil \mathrm{M_z} &= \mathrm{N_z^{-1}} \sum_{\mathrm{i} \in \mathrm{z}} (\mathrm{X_i} - \mathrm{X}) \cdot I[\mathrm{X_i} > \mathrm{X}] \end{aligned} \tag{76}$$

Applying (63),

$$\lceil \mathrm{M_z} = \mathrm{N_z^{-1}} \sum_{\mathrm{o}} \Big\{ \sum_{\mathrm{i} \in \mathrm{o,z}} (\mathrm{X_i} - \mathrm{X}) \cdot I[\mathrm{X_i} > \mathrm{X}] \Big\} = \mathrm{N_z^{-1}} \sum_{\mathrm{o}} \Big\{ (\lceil \mathrm{X_{o,z}} - \mathrm{X}) \cdot \lceil \mathrm{N_{o,z}} \Big\} \tag{77}$$

$\lceil \mathrm{X_{o,z}}$ is defined as the mean value of $\mathrm{X_i}$ in z restricted both to category o, and to the condition of being greater than the central point X; and $\lceil \mathrm{N_{o,z}}$ is the count satisfying the same. $\lfloor \mathrm{X_{o,z}}$ and $\lfloor \mathrm{N_{o,z}}$ are similar, for the condition of being less than the central point.

$$\lceil \mathrm{M_z} = \sum_{\mathrm{o}} \overbrace{\Big(\lceil \mathrm{X_{o,z}} - \mathrm{X} \Big)}^{\text{Shift}} \cdot \overbrace{\lceil \mathrm{N_{o,z}} / \mathrm{N_z}}^{\text{Share}} \;,\; \lfloor \mathrm{M_z} = \sum_{\mathrm{o}} \overbrace{\Big(\lfloor \mathrm{X_{o,z}} - \mathrm{X} \Big)}^{\text{Shift}} \cdot \overbrace{\lfloor \mathrm{N_{o,z}} / \mathrm{N_z}}^{\text{Share}} \tag{78}$$

The instrument follows from delocalizing the Shifts over z and lagging the Shares,

$$
\begin{aligned}
\lceil \tilde{\mathrm{M}}_{\mathrm{t,z}} &= \sum_{\mathrm{o}} \overbrace{\Big(\lceil \mathrm{X}_{\mathrm{o,t}} - \mathrm{X}_{\mathrm{t}}\Big)}^{\text{Shift}} \cdot \overbrace{\lceil \mathrm{N}_{\mathrm{o,t}}/\mathrm{N}_{\mathrm{t}}}^{\text{Share}} \\
\lfloor \tilde{\mathrm{M}}_{\mathrm{t,z}} &= \sum_{\mathrm{o}} \overbrace{\Big(\lfloor \mathrm{X}_{\mathrm{o,t}} - \mathrm{X}_{\mathrm{t}}\Big)}^{\text{Shift}} \cdot \overbrace{\lfloor \mathrm{N}_{\mathrm{o,t}}/\mathrm{N}_{\mathrm{t}}}^{\text{Share}}
\end{aligned} \tag{79}
$$

$$\tilde{\mathrm{M}}_{\mathrm{t,z}} = \lfloor \tilde{\mathrm{M}}_{\mathrm{t,z}} + \lceil \tilde{\mathrm{M}}_{\mathrm{t,z}}$$

Although $\lceil \mathrm{X}_{\mathrm{o,t}}$, $\lceil \mathrm{N}_{\mathrm{o,t}}$, $\lfloor \mathrm{X}_{\mathrm{o,t}}$, $\lfloor \mathrm{N}_{\mathrm{o,t}}$ are unusual mathematical objects, they follow directly from the accounting identity as the natural components for the instrument, are straightforward to construct from national Industry o data.

3.6.3 Bin Shares and Gini

Bin Shares are objects that can flexibly map the shape of a distribution. Bin Shares are,

$$\mathrm{P}_{\mathrm{h,z}} = \mathrm{N}_{\mathrm{h,z}}/\mathrm{N}_{\mathrm{z}}$$

where $\mathrm{N}_{\mathrm{h,z}}$ is the count of persons i in Locality z who fall into bin h; and for i to fall into bin h means that the variable X_{i} falls between an upper and lower cutoff particular to h. Table 3.1 provides a concrete example of Bin Shares. Here, the variable X_{i} is each individual i's self reported yearly earnings, Y_{i}, and the bin cutoffs are defined by deciles of the national earnings distribution.

The right-most column of Table 3.1 reports the Bin Shares for z = *Boston* (the Commuting Zone) in 2010. If the earnings distribution in Boston were a perfect match to the national earnings distribution, then all of Boston's Bin Shares would be equal to 0.10, because the bins are deciles of the national distribution. Instead, because Boston is wealthy, it has above par shares in the three highest bins, especially the top bin. The Cutoff column gives the lower cutoff for each national decile bin h. Individual earnings are reported to the nearest $1,000, yielding clean bin cutoffs. $\mathrm{Y}_{\mathrm{h,t}}$ is the national average within each bin, which may serve as a (discrete approximation) representative value for any Y_i within each.

Table 3.1: (Earnings Decile) Bin Shares for Boston in 2010

Decile (h)	Cutoff	$Y_{h,t}$ ($\$10^3$)	$P_{h,\ t=\mathit{2010},\ z=\mathit{Boston}}$
1	0	0	0.0970
2	6.200	3.176	0.0859
3	13.00	9.933	0.0764
4	20.00	17.12	0.0816
5	26.00	23.63	0.0589
6	34.00	30.16	0.0835
7	42.00	38.23	0.0938
8	55.00	48.90	0.1157
9	76.00	65.16	0.1257
10	-	130.8	0.1814

Bin shares $P_{h,t,z} = N_{h,t,z}/N_{t,z}$ map the shape of a distribution. Here, bin cutoffs are defined by deciles of the national earnings distribution; the Cutoff column gives the lower cutoff of each. $Y_{h,t}$ is the national average within each bin. Data are from the American Community Survey via IPUMS USA.

Boustan et al. (2013) use Bin Shares as the Shares in their instrument for the Gini coefficient of income. However, as Bin Shares are simply an approximate mapping of the local income distribution, they are similar in terms of exogeneity to the average (or any other summary) of the local income distribution, which would often be considered endogenous even in the lag. As an alternative, I the cast the Bin Shares themselves as endogenous variables, for which we can construct classical Shift-Share instruments. Applying the Bin Share, $P_{h,t,z} = N_{h,t,z}/N_{t,z}$, as $X_{t,z}$ in the Classical Accounting

Identity, we have as an accounting identity,

$$\mathrm{P}_{\mathrm{h,t,z}} = \sum_{\mathrm{o}} \overbrace{\mathrm{N}_{\mathrm{h,o,t,z}} \Big/ \mathrm{N}_{\mathrm{o,t,z}}}^{\text{Shift}} \cdot \overbrace{\mathrm{N}_{\mathrm{o,t,z}} \Big/ \mathrm{N}_{\mathrm{t,z}}}^{\text{Share}} \tag{80}$$

The Shift-Share instrument for the Bin Share then follows from delocalizing the Shift vector over Localities z, and lagging the Industrial Share vector.

$$\tilde{\mathrm{P}}_{\mathrm{h,t,z}} = \sum_{\mathrm{o}} \overbrace{\mathrm{N}_{\mathrm{h,o,t}} \Big/ \mathrm{N}_{\mathrm{o,t}}}^{\text{Shift}} \cdot \overbrace{\mathrm{N}_{\mathrm{o},t,\mathrm{z}} \Big/ \mathrm{N}_{t,\mathrm{z}}}^{\text{Share}} \tag{81}$$

These can then be applied in the place of the endogenous Bin Shares in the discretized Gini, yielding a Shift-Share instrument for the Gini:

$$\frac{\sum_{h} \sum_{\mathrm{h}} \overbrace{\tilde{\mathrm{P}}_{\mathrm{h,t,z}} \tilde{\mathrm{P}}_{h,\mathrm{t,z}}}^{\text{Shift}\cdot\text{Share}} \cdot \overbrace{\left| \mathrm{Y}_{\mathrm{h,t}} - \mathrm{Y}_{h,\mathrm{t}} \right|}^{\text{Push}}}{2 \sum_{h} \underbrace{\tilde{\mathrm{P}}_{h,\mathrm{t,z}}}_{\text{Shift}\cdot\text{Share}} \cdot \underbrace{\mathrm{Y}_{h,\mathrm{t}}}_{\text{Push}}} \tag{82}$$

Although Boustan et al. (2013)'s instrument may be validly exogenous, (82) provides arguably stronger guarantees of exogeneity, more like that of a classical Shift-Share instrument. The upcoming section tests multiple related instruments for the Gini in practice.

3.7 Application: Inequality and Single Parenting

There are many reasons that we might expect earnings inequality to result in higher rates of single parenting. However, there has been little direct work on this topic in economics, despite its importance.[109] Gould and Paserman (2003) articulate

[109] Chetty et al. (2014) find single parenting to be the strongest and most robust single predictor of socioeconomic immobility.

a theory as to why inequality would result in lower rates of marriage: In more unequal markets, women hold out longer in hopes of securing more desirable matches. Kearney and Levine (2014) discuss several other theories pertaining to single parenting directly, emphasizing rather hope*less*ness: Underprivileged women in more unequal markets may have little hope of securing decent marriage matches (or decent careers for themselves), anyway. Therefore, the burden of single parenting makes relatively little difference to their life prospects.

In examining the relationship between inequality and single parenting propensity (or marriage rates), there is often an elephant in the room. Although both Gould and Paserman (2003) and Kearney and Levine (2014) use rich sets of control variables to test for alternative explanatory factors, neither clearly addresses the question of direct reverse-causal effects of single parenting (or marriage rates) on inequality. Indeed, Chetty et al. (2014) find single parenting to be the strongest and most robust single predictor of socioeconomic immobility: Although not causal, this is highly suggestive. Single parenting may exacerbate inequality, either directly - by distorting women's career prospects, or distorting men's career incentives - or via societal trauma more generally.

As an application in this paper, I test various alternative instruments for the Gini coefficient of earnings to estimate the effect of earnings inequality on rates of single parenting. The Gini is one of many plausible (but perhaps the single most well-known) measures of inequality. Because my goal is only to examine the implications of using different instruments, I use a (causal) "reduced-form" approach. That is, rather than deriving a model of equilibrium earnings and single parenting from assumed utility functions (i.e. a "structural" approach), I begin rather from the more downstream assumption that single parenting's equilibrium response to earnings can be written in a general log-linear form,

$$\mathit{log}\ (\ \mathrm{P}_{\mathrm{t,z}}^{\mathit{SingleParent}}\ )\ =\ \beta^{\mathit{Inequality}}\ \cdot\ \mathit{log}\ (\ \mathit{gini}\ [\ \mathrm{Y}_{\mathrm{i}\ \in\ \mathrm{t,z}}\ ]\ )\ +\ \{\ \beta^{x}\ \cdot\ \mathrm{x}_{\mathrm{t,z}}\ \} \tag{83}$$

where $\mathrm{P}_{\mathrm{t,z}}^{SingleParent}$ is the fraction of adults in Commuting Zone z who are single parents, and $\mathrm{x}_{\mathrm{t,z}}$ are controls. $\beta^{Inequality}$ is the causal elasticity of single parenting with respect to inequality: It encompasses the *sum* effect of all mechanisms through which inequality would motivate single parenting - including both the "hope" of Gould and Paserman (2003), and the "hopelessness" of Kearney and Levine (2014).

Although equation (83) represents all effects of inequality on single parenting propensity, it does not represent any reverse effects - that is, of single parenting on inequality.[110] For this reason, an ordinary regression estimation of (83) would be confounded by simultaneity bias, and fail to recover $\beta^{Inequality}$. The purpose of instrumenting is to restrict the information about the Gini that enters into the estimation of (83) to only variation that cannot be driven by single parenting. This removes any reverse-casual effects, hence identifying $\beta^{Inequality}$.

As discussed throughout this paper, Shift-Share instruments use delocalizations over space and time. These delocalizations can make the instruments causally prior (exogenous) with respect to almost any outcome. The Shift-Share instrument restricts to information from previous time periods (Shares), and from other cross sectional units (Localities) beyond z - typically the national or global average over all such units. Therefore, the instrument is valid so long as the outcome cannot affect either Shares in previous time periods, or national averages beyond z in the current time period. These conditions can usually be supposed to hold, given that Shares are chosen to represent deep characteristics of Localities z.

Like Shift-Share instruments, the instrument of Boustan et al. (2013), which I call the Lagged Bin Share Gini, can be considered valid in a wide variety of settings. In Boustan et al. (2013), the outcome of interest is local government taxation and expenditure. Enamorado et al. (2016) use the same instrument, but for a very different outcome variable - violent crime rates. In either case, the explanatory variable of interest is the local Gini coefficient of income, as a measure of inequality. Although

[110] The sum of these reverse effects would follow an analogous, additional equation (unwritten).

the reason for inequality to drive taxation is unrelated to the reason for inequality to drive crime, the instrument is valid for the same reason in either setting: *present* values of the outcome variable cannot drive *past* values of the Shares that are used to construct the instrument.

I apply the Lagged Bin Share Gini (and my alternative Shift-Share Gini) in yet another setting, that is, with single parenting rates as the outcome variable. As in the settings of local taxation and of crime rates, the instrument is valid for the same reason: *present* period single parenting rates cannot drive *past* values of the Shares. This raises the question: What advantage do Shift-Share instruments have over simply using lags as instruments? That is, why not use lagged values of the Gini itself as instrument for the Gini in the present?

Indeed, it is generally plausible to assume that present values of *any* variable do not affect past values of any other variable, and as such, the use of simple lags as instruments is widespread. The exception would be if agents are "forward looking" (able to predict the future), and also able to adjust the lagged explanatory variable accordingly (in response to the future). The advantage of Shift-Share instruments over simple lag instruments comes from the split between Shift and Share: Only the Shares are lagged, while the Shifts are delocalized in a stronger way. Shares are chosen to represent deep characteristics - things that cannot be readily adjusted in response to future expectations. The remaining, shallower or more adjustable components of the explanatory variable can thus be resolved into the Shift, and hence neutralized by the stronger delocalization.

I argue that the Lagged Bin Share Gini is essentially similar to a simple lag instrument. Although a lag in itself may yield a valid instrument, a true Shift-Share instrument takes the additional step of more strongly delocalizing a component (the Shift) of the explanatory variable that is most vulnerable to endogeneity. The Lagged Share Gini instrument (with a Frozen Lag t) is,

$$\tilde{\mathrm{X}}_{\mathrm{t,z}} = \frac{\sum_h \sum_{\mathrm{h}} \overbrace{\mathrm{P}_{\mathrm{h},t,\mathrm{z}} \mathrm{P}_{h,t,\mathrm{z}}}^{\text{Share}} \cdot \overbrace{\left| \mathrm{Y}_{\mathrm{h,t}} - \mathrm{Y}_{h,\mathrm{t}} \right|}^{\text{Push}}}{2 \sum_h \underbrace{\mathrm{P}_{h,t,\mathrm{z}}}_{\text{Share}} \cdot \underbrace{\mathrm{Y}_{h,\mathrm{t}}}_{\text{Push}}} \tag{84}$$

With Updating Lags rather,[111]

$$\tilde{\mathrm{X}}_{\mathrm{t,z}} = \frac{\sum_h \sum_{\mathrm{h}} \overbrace{\mathrm{P}_{\mathrm{h,t-1,z}} \mathrm{P}_{h,\mathrm{t-1,z}}}^{\text{Share}} \cdot \overbrace{\left| \mathrm{Y}_{\mathrm{h,t}} - \mathrm{Y}_{h,\mathrm{t}} \right|}^{\text{Push}}}{2 \sum_h \underbrace{\mathrm{P}_{h,\mathrm{t-1,z}}}_{\text{Share}} \cdot \underbrace{\mathrm{Y}_{h,\mathrm{t}}}_{\text{Push}}} \tag{85}$$

For comparison, a simple lag of the Gini is,

$$\mathrm{X}_{\mathrm{t-1,z}} = \frac{\sum_h \sum_{\mathrm{h}} \overbrace{\mathrm{P}_{\mathrm{h,t-1,z}} \mathrm{P}_{h,\mathrm{t-1,z}}}^{\text{Share}} \cdot \overbrace{\left| \mathrm{Y}_{\mathrm{h,t-1}} - \mathrm{Y}_{h,\mathrm{t-1}} \right|}^{\text{Push}}}{2 \sum_h \underbrace{\mathrm{P}_{h,\mathrm{t-1,z}}}_{\text{Share}} \cdot \underbrace{\mathrm{Y}_{h,\mathrm{t-1}}}_{\text{Push}}} \tag{86}$$

The difference between (85) and (86) is that the latter lags the Push, while the former uses present values for the Push. However, because the Push is invariant over Localities z in either case, there is little scope for this change to yield meaningful differences in terms of exogeneity.[112] In other words, because the Shares in this case contain essentially all of the information in the original Gini that that is distinctive per Locality z, lagging these Shares is similar to lagging the whole Gini. By contrast, a Shift-Share instrument would sequester some of the distinctive information - that

[111]See Section 3.4.1. Frozen Lags have the clear downside that they make the instrument more relevant in earlier time periods (closer to the lag), and weaker as time goes on. Updating Lags have no clear downside in my opinion.

[112]The Lagged Bin Share instrument may still have an advantage over the simple lag in terms of relevance, although it does not in terms of exogeneity. In other words, the Lagged Bin Share Gini may be better correlated with the present Gini than the lagged Gini is.

which is most sensitive to endogeneity - into a Shift vector which is delocalized, hence yielding a stronger claim of exogeneity than that offered just by a lag.

Whether a lag in itself is sufficient to yield exogeneity depends on context - particularly, how plausible it is that the explanatory variable responds to *future* values of the outcome of interest. In the setting of Boustan et al. (2013), this question would be whether the local income distribution responds to future changes in local taxation and expenditure policy. Lower earning people for example, if accurately predicting future policy, may select into localities in which future taxation and expenditure are on an upward trajectory. With single parenting rates as the outcome of interest, an analogous confounding scenario may be that wealthier people avoid localities in which single parent homes are on the rise.

Where a lag instrument relies on the explanatory variable to not respond to future changes in the outcome variable, a Shift-Share instrument rather relies on only the Shares to not do so. To accomplish this purpose, it is vital that Shares represent deep characteristics of cross sectional units (Localities), such as can be viewed as causal first movers of the economic system under study. The industrial profile each Locality (the classical Share vector) meets this purpose because it arises from geographical and historical forces that are beyond the scope of the model. For example, shipping relies on access to water, and locations are bound to particular industries, such as Detroit to manufacturing, for historical reasons that cannot be easily adjusted.

Unlike classical industrial Shares, the Shares of the Lagged Bin Share instrument needn't coincide with anything particularly fundamental to localities. These Shares are Bin Shares, which simply map the local income distribution. They are sensitive to anything that alters the local income distribution, including selective migration of higher or lower earning individuals. As such, these Shares are more vulnerable to confoundedness by future expectations than Shares of a typical Shift-Share instrument would be. This is the flip side of lacking a Shift component that is delocalized over the cross section. The local averages that are most vulnerable to endogeneity would be in the Shift rather than the Share, and hence neutralized by the Shift delocalization.

Where the Lagged Bin Share Gini replaces the endogenous Bin Share with a lagged Bin Share, my Shift-Share Gini rather replaces it with a classical Shift-Share instrument for the Bin Share itself. Although it is its own kind of share, the Bin Share does not constitute the Share of a Shift-Share instrument, for two reasons that are flip sides of the same coin. First, there is no corresponding Shift - that is, a component distinctive to each Locality that is delocalized in the instrument. Second, indeed because there is no Shift component to carry the regressor's local information that is most vulnerable to endogeneity, the Bin Share (even lagged) is left vulnerable. My Shift-Share Gini fixes this problem by resolving the Bin Share into an underlying Shift and exogenous Share.

To examine concrete implications of each of the above versions of the Gini instrument, I estimate (86) using each.[113] OLS estimation of (86) should be confounded by reverse causality in so far as single parenting rates have effects on the distribution of earnings. If both the forward and reverse effects are positive - that is, if inequality drives more single parenting, and single parenting also drives more inequality - then it is natural for the OLS estimate to be biased downward in magnitude.[114] Indeed, Table 3.2 shows that the elasticity estimate reported by OLS is smaller than that of 2SLS using any of the instrument alternatives.

[113]Source data are from the US Census and American Community Survey via IPUMS USA, Ruggles et al. (2020). I aggregate by Commuting Zones (CZs) as Localities z. CZs are defined based on actual commuting patterns in 1990, and hence capture local labor markets; see Tolbert and Sizer (1996) and Autor and Dorn (2013).

[114]Although a mutually positive reverse-causal effect will increase the correlation between the two variables, it will nonetheless decrease the magnitude of the OLS slope coefficient estimate, by increasing the "run" in "slope = rise/run" more than it increases the "rise."

Table 3.2: Elasticity of Single Parenting w.r.t. Gini of Earnings

	OLS	2SLS			
	Gini	Lagged Share Gini	Lagged Gini	Delocalized Shift Gini	Shift-Share Gini
Gini	0.654***	1.077***	1.105***	1.184***	1.487***
	(0.114)	(0.271)	(0.268)	(0.221)	(0.369)
Gini	0.465***	1.002***	0.997***	1.081***	1.217***
	(0.117)	(0.272)	(0.271)	(0.227)	(0.401)
Mean	-0.280***	-0.226***	-0.209**	-0.165**	-0.151*
	(0.045)	(0.082)	(0.082)	(0.083)	(0.088)

Observations are 722 Commuting Zones by 5 time periods 1980-2019. Dependent variable is log single parenting rate; regressor is log Gini of earnings. Controls are region by year effects, as well log average earnings in the lower panel. The 2SLS columns each use a different instrument for the Gini as stated, and also instrument average earnings with the Shift-Share average.

The results given in Table 3.2 illustrate two main points concerning the different instruments. First is that the Lagged Bin Share instrument (Column 2) gives numerically very similar results as the simple lag instrument (Column 3). This is an empirical confirmation of my theoretical argument that the Lagged Bin Share instrument is similar to the simple lag in terms of exogeneity. The second main point is that - supposing the Shift-Share instrument (Column 5) is fully exogenous, and hence yielding the most accurate result - then the Lagged Bin Share and lag instruments are getting part of the way there, that is, partially but not fully correcting the bias from OLS.[115] This is indeed what we ought to expect given that the Shift-Share instrument uses both a (Share) lag and a Shift delocalization, and that both of these are essential in correcting bias.

[115]Likewise, the Delocalized Shift instrument (Column 4) corrects bias only part way. This uses the original (non-lagged) Share vector in conjunction with the delocalized Shift vector.

3.8 Conclusion

This paper may be unique in focusing on the generative process by which a researcher may arrive at Shift-Share instruments. By deriving the instrument directly from an accounting identity of the explanatory variable, one can appreciate both the essential features of Shift-Share instruments, and the scope of their potential varieties. Using my simple approach for understanding Shift-Share instruments as modified accounting identities, I closely compare a wide variety of instruments from the literature. I then also develop general formulas for several new varieties - instruments for variances, skews, mean absolute deviations, and Gini coefficients.

As an empirical application, I measure the effect of earnings inequality on rates of single parenting, using multiple alternative instruments for the Gini coefficient of earnings. The empirical results illustrate core themes from earlier in the paper. That is, Shift-Share instruments both delocalize (replace with nonlocal averages) the more endogenous part of the accounting identity (the Shift vector), and lag the more exogenous part of the accounting identity (the Share vector). Empirically, I show that instruments that do only one or the other - delocalize the Shift, or lag the Share - also correct bias, but only part way. Thus, although each of these steps - delocalizing the Shift, and lagging the Share - provides its own argument of exogeneity, both make meaningful contributions to the exogeneity of the Shift-Share instrument as a whole.

References

Abraham, Martin, Sebsatian Bähr, and Mark Trappmann. 2019. "Gender Differences in Willingness to Move for Interregional Job Offers." *Demographic Research*, 40(53): 1537–1602.

Adão, Rodrigo, Michal Kolesár, and Eduardo Morales. 2019. "Shift-Share Designs: Theory and Inference." *Quarterly Journal of Economics*, 134(4): 1949–2010.

Ales, Laurence, Musab Kurnaz, and Christopher Sleet. 2015. "Technical change, wage inequality, and taxes." *American Economic Review*, 105(10): 3061–3101.

Altig, David, and Charles T Carlstrom. 1999. "Marginal tax rates and income inequality in a life-cycle model." *American Economic Review*, 89(5): 1197–1215.

Angrist, Josh. 2002. "How Do Sex Ratios Affect Marriage and Labor Markets? Evidence From America's Second Generation." *Quarterly Journal of Economics*, 117(3): 997–1038.

Attanasio, Orazio P, and Katja M Kaufmann. 2017. "Education Choices and Returns on the Labor and Marriage Markets: Evidence from Data on Subjective Expectations." *Journal of Economic Behavior & Organization*, 140: 35–55.

Autor, David. 2019. "Work of the Past, Work of the Future." *AEA Papers and Proceedings*, 109: 1–32.

Autor, David, and David Dorn. 2013. "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market." *American Economic Review*, 103(5): 1553–1597.

Autor, David, David Dorn, and Gordon Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review*, 103(6): 2121–2168.

Autor, David, David Dorn, and Gordon Hanson. 2019. "When Work Disappears: Manufacturing Decline and the Falling Marriage Market Value of Young Men." *American Economic Review*, 1(2): 161–178.

Banerjee, Abhijit V, and Sendhil Mullainathan. 2008. "Limited attention and income distribution." *American Economic Review*, 98(2): 489–93.

Bartik, Timothy. 1991. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Basu, Rounaq, and Joseph Ferreira. 2021. "Sustainable mobility in auto-dominated Metro Boston: Challenges and opportunities post-COVID-19." *Transport Policy*, 103: 197–210.

Baumer, Eric P, and Janet L Lauritsen. 2010. "Reporting crime to the police, 1973–2005: a multivariate analysis of long-term trends in the National Crime Survey (NCS) and National Crime Victimization Survey (NCVS)." *Criminology*, 48(1): 131–185.

Baum-Snow, Nathaniel, and Fernando Ferreira. 2015. "Causal Inference in Urban and Regional Economics." *Handbook of Regional and Urban Economics*, 5: 3–68.

Bayer, Patrick, Robert McMillan, Alvin Murphy, and Christopher Timmins. 2016. "A Dynamic Model of Demand for Houses and Neighborhoods." *Econometrica*, 84(3): 893–942.

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, 76(2): 169–217.

Becker, Gary S. 1973. "A Theory of Marriage: Part I." *Journal of Political Economy*, 81(4): 813–846.

Bertrand, Marianne, Jessica Pan, and Emir Kamenica. 2013. "Gender Identity and Relative Income Within Households." *Quarterly Journal of Economics*, 130(2): 571–614.

Bhargava, Saurabh, and Dayanand Manoli. 2015. "Psychological frictions and the incomplete take-up of social benefits: Evidence from an IRS field experiment." *American Economic Review*, 105(11): 3489–3529.

Blackburn, McKinley L. 2010. "Internal Migration and the Earnings of Married Couples in the United States." *Journal of Economic Geography*, 10(1): 87–111.

Blanchard, Olivier, and Lawrence Katz. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity*, 23(1): 1–76.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2022. "Quasi-Experimental Shift-Share Research Designs." *The Review of Economic Studies*, 89(1): 181–213.

Bound, John, and Harry Holzer. 2000. "Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s." *Journal of Labor Economics*, 18(1): 20–54.

Boustan, Leah, Fernando V Ferreira, Hernan Winkler, and Eric M Zolt. 2013. "The Effect of Rising Income Inequality on Taxation and Public Expenditures: Evidence from U.S. Municipalities and School Districts, 1970-2000." *The Review of Economics and Statistics*, 95(1291).

Brandén, Maria, and Karen Haandrikman. 2019. "Who Moves to Whom? Gender Differences in the Distance Moved to a Shared Residence." *European Journal of Population*, 35: 435–458.

Brush, Jesse. 2007. "Does income inequality lead to more crime? A comparison of cross-sectional and time-series analyses of United States counties." *Economics letters*, 96(2): 264–268.

Burdett, Kenneth, Ricardo Lagos, and Randall Wright. 2003. "Crime, inequality, and unemployment." *American Economic Review*, 93(5): 1764–1777.

Cameron, Lisa, Xin Meng, and Dandan Zhang. 2019. "China's Sex Ratio and Crime: Behavioural Change or Financial Necessity?" *The Economic Journal*, 129(618): 790–820.

Card, David. 2001. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics*, 19(1): 22–64.

Carvalho, Leandro S, Stephan Meier, and Stephanie W Wang. 2016. "Poverty and economic decision-making: Evidence from changes in financial resources at payday." *American economic review*, 106(2): 260–84.

Chalfin, Aaron. 2015. "The long-run effect of Mexican immigration on crime in US cities: Evidence from variation in Mexican fertility rates." *American Economic Review*, 105(5): 220–25.

Charles, Kerwin Kofi, and Ming Ching Luoh. 2010. "Male Incarceration, the Marriage Market, and Female Outcomes." *The Economic Journal*, 92(3): 614–627.

Chetty, Raj, Nathaniel Hendren, and Lawrence F Katz. 2016. "The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment." *American Economic Review*, 106(4): 855–902.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *Quarterly Journal of Economics*, 129(4): 1553–1623.

Chiappori, Pierre-André, Bernard Salanié, and Yoram Weiss. 2017. "Partner Choice, Investment in Children, and the Marital College Premium." *American Economic Review*, 107(8): 2109–2167.

Choe, Jongmook. 2008. "Income inequality and crime in the United States." *Economics Letters*, 101(1): 31–33.

Coleman, Clive, and Jenny Moynihan. 1996. *Understanding crime data: Haunted by the dark figure.* Vol. 120, Open University Press Buckingham.

Compton, Janice, and Robert A Pollak. 2007. "Why Are Power Couples Increasingly Concentrated in Large Metropolitan Areas?" *Journal of Labor Economics*, 25(3): 475–512.

Compton, Janice, and Robert A Pollak. 2014. "Family Proximity, Childcare, and Women's Labor Force Attachment." *Journal of Urban Economics*, 79: 72–90.

Croson, Rachel, and Uri Gneezy. 2009. "Gender Differences in Preferences." *Journal of Labor Economics*, 47(2): 448–474.

Dancygier, Rafaela, Naoki Egami, Amaney Jamal, and Ramona Rischke. 2021. "Hate Crimes and Gender Imbalances: Fears over Mate Competition and Violence against Refugees." *American Journal of Political Science*, 66(2): 501–515.

Davis, Jonathan MV, and Sara B Heller. 2020. "Rethinking the benefits of youth employment programs: The heterogeneous effects of summer jobs." *Review of Economics and Statistics*, 102(4): 664–677.

Demombynes, Gabriel, and Berk Özler. 2005. "Crime and local inequality in South Africa." *Journal of development Economics*, 76(2): 265–292.

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000." *American Economic Review*, 106(3): 479–524.

Dollar, Cindy Brooks. 2017. "Sex Ratios, Single Motherhood, and Gendered Structural Relations: Examining Female-Headed Families Across Racial-Ethnic Populations." *Sociological Focus*, 50(4): 375–390.

Donnelly, Kristin, Jean M Twenge, and Malissa A Clark. 2016. "Attitudes Toward Women's Work and Family Roles in the United States, 1976-2013." *Psychology of Women Quarterly*, 40(1): 41–54.

Dorn, David. 2009. "Essays on Inequality, Spatial Interaction, and the Demand for Skills." PhD diss. University of St. Gallen.

Duquennois, Claire. 2022. "Fictional money, real costs: Impacts of financial salience on disadvantaged students." *American Economic Review*, 112(3): 798–826.

Edlund, Lena. 2005. "Sex and the City." *Scandinavian Journal of Economics*, 107(1): 25–44.

Edlund, Lena, Hongbin Li, Junjian Yi, and Junsen Zhang. 2013. "Sex Ratios and Crime: Evidence from China." *The Review of Economics and Statistics*, 95(5): 1520–1534.

Ehrlich, Isaac. 1973. "Participation in illegitimate activities: A theoretical and empirical investigation." *Journal of political Economy*, 81(3): 521–565.

Ehrlich, Isaac. 1975. "The Deterrent Effect of Capital Punishment: A Question of Life and Death." *American Economic Review*, 65(3): 397–417.

Ehrlich, Isaac. 1977. "Capital punishment and deterrence: Some further thoughts and additional evidence." *Journal of Political Economy*, 85(4): 741–788.

Ehrlich, Issac. 1970. "Participation in Illegitimate Activities: An Economic Analysis." PhD diss. Columbia University.

Enamorado, Ted, Luis F López-Calva, Carlos Rodríguez-Castelán, and Hernán Winkler. 2016. "Income inequality and violent crime: Evidence from Mexico's drug war." *Journal of Development Economics*, 120: 128–143.

Engelhardt, Bryan, Guillaume Rocheteau, and Peter Rupert. 2008. "Crime and the labor market: A search model with optimal contracts." *Journal of Public Economics*, 92(10-11): 1876–1891.

Fajnzylber, Pablo, Daniel Lederman, and Norman Loayza. 2002. "Inequality and violent crime." *The journal of Law and Economics*, 45(1): 1–39.

Farhi, Emmanuel, and Xavier Gabaix. 2020. "Optimal taxation with behavioral agents." *American Economic Review*, 110(1): 298–336.

Federal Bureau of Investigation. 1980. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 1977-1983." United States Department of Labor. `https://www.icpsr.umich.edu/web/NACJD/studies/8703` (accessed August 13, 2021).

Federal Bureau of Investigation. 1990. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 1990." United States Department of Labor. `https://www.icpsr.umich.edu/web/NACJD/studies/9785` (accessed August 12, 2021).

Federal Bureau of Investigation. 2000. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2000." United States Department of Labor. `https://www.icpsr.umich.edu/web/NACJD/studies/3451` (accessed August 12, 2021).

Federal Bureau of Investigation. 2010. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2010." United States Department of Labor. `https://www.icpsr.umich.edu/web/NACJD/studies/33526` (accessed August 12, 2021).

Federal Bureau of Investigation. 2016. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2016." United States Department of Labor. https://www.icpsr.umich.edu/web/NACJD/studies/37059 (accessed August 13, 2021).

Federal Bureau of Investigation. 2018*a*. "*UCR Property Crime Definition.*" https://ucr.fbi.gov/crime-in-the-u.s/2018/crime-in-the-u.s.-2018/topic-pages/property-crime (Accessed: September 28, 2021).

Federal Bureau of Investigation. 2018*b*. "*UCR Violent Crime Definition.*" https://ucr.fbi.gov/crime-in-the-u.s/2018/crime-in-the-u.s.-2018/topic-pages/violent-crime (Accessed: September 28, 2021).

Fisman, Raymond, Sheena S Iyengar, Emir Kamenica, and Itamar Simonson. 2006. "Gender Differences in Mate Selection: Evidence From a Speed Dating Experiment." *Quarterly Journal of Economics*, 121(2): 673–697.

Florida, Richard, and Charlotta Mellander. 2016. "Rise of the startup city: The changing geography of the venture capital financed innovation." *California Management Review*, 59(1): 14–38.

Gautier, Pieter A, Michael Svarer, and Coen N Teulings. 2010. "Marriage and the City: Search Frictions and Sorting of Singles." *Journal of Urban Economics*, 67(2): 206–218.

Glaeser, Edward L, and David C Maré. 2001. "Cities and Skills." *Journal of Labor Economics*, 19(2): 316–342.

Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review*, 104(4): 1091–1119.

Goldin, Claudia, Lawrence F Katz, and Ilyana Kuziemko. 2006. "The Homecoming of American College Women: The Reversal of the College Gender Gap." *The Journal of Economic Perspectives*, 20(4): 133–156.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. "Bartik Instruments: What, When, Why, and How." *American Economic Review*, 110(8): 2586–2624.

Gould, Eric D, and Daniele M Paserman. 2003. "Waiting for Mr. Right: Rising Inequality and Declining Marriage Rates." *Journal of Urban Economics*, 53(2): 257–281.

Gould, Eric D, Bruce A Weinberg, and David B Mustard. 2002. "Crime rates and local labor market opportunities in the United States: 1979–1997." *Review of Economics and statistics*, 84(1): 45–61.

Greenwood, Michael J. 1997. "Internal Migration in Developed Countries." *Handbook of Population and Family Economics*, 1: 647–720.

Gubhaju, Bina, and Gordon F De Jong. 2009. "Individual versus Household Migration Decision Rules: Gender and Marital Status Differences in Intentions to Migrate in South Africa." *International Migration*, 47(1): 31–61.

Harknett, Kristen. 2008. "Mate Availability and Unmarried Parent Relationships." *Demography*, 45(3): 555–571.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L Violante. 2017. "Optimal Tax Progressivity: An Analytical Framework." *Quarterly Journal of Economics*, 132(4): 1693–1754.

Hipp, John R. 2007. "Income inequality, race, and place: Does the distribution of race and class within neighborhoods affect crime rates?" *Criminology*, 45(3): 665–697.

Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. 2016. "Long-run impacts of childhood access to the safety net." *American Economic Review*, 106(4): 903–34.

International Centre for Prison Studies. 2017-2018. "World Prison Brief Data." `https://www.prisonstudies.org/world-prison-brief-data` (accessed February 28, 2022).

Jang, Bohyun, John Casterline, and Anastasia Snyder. 2014. "Migration and Marriage: Modeling the Joint Process." *Demographic Research*, 30(47): 1339–1366.

Kang, Songman. 2016. "Inequality and crime revisited: Effects of local inequality and economic segregation on crime." *Journal of Population Economics*, 29(2): 593–626.

Kaur, Supreet, Sendhil Mullainathan, Suanna Oh, and Frank Schilbach. 2021. "Do Financial Concerns Make Workers Less Productive?" National Bureau of Economic Research.

Kearney, Melissa S, and Phillip B Levine. 2014. "Income Inequality and Early Nonmarital Childbearing." *The Journal of Human Resources*, 49(1): 1–31.

Kelly, Morgan. 2000. "Inequality and crime." *Review of economics and Statistics*, 82(4): 530–539.

Kennan, John, and James R Walker. 2011. "The Effect of Expected Income on Individual Migration Decisions." *Econometrica*, 79(1): 211–251.

Kennedy, Leslie W. 1988. "Going it alone: Unreported crime and individual self-help." *Journal of Criminal Justice*, 16(5): 403–412.

Kessler, Judd B, Sarah Tahamont, Alexander M Gelber, and Adam Isen. 2021. "The Effects of Youth Employment on Crime: Evidence from New York City Lotteries." National Bureau of Economic Research.

Kröhnert, Steffen, and Sebastian Vollmer. 2012. "Gender-Specific Migration from Eastern to Western Germany: Where Have All the Young Women Gone?" *International Migration*, 50(5): 95–112.

Levitt, Steven D. 1998. "The relationship between crime reporting and police: Implications for the use of uniform crime reports." *Journal of quantitative criminology*, 14(1): 61–81.

Løken, Katrine V, Kjell Erik Lommerud, and Shelly Lundberg. 2013. "Your Place or Mine? On the Residence Choice of Young Couples in Norway." *Demography*, 50(1): 285–310.

Matarazzo, Thomas, Mohammad Vazifeh, Shamim Pakzad, Paolo Santi, and Carlo Ratti. 2017. "Smartphone data streams for bridge health monitoring." *Procedia engineering*, 199: 966–971.

McFadden, Daniel. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior." *Frontiers in Econometrics*, 105–145.

Modestino, Alicia Sasser. 2019. "How do summer youth employment programs improve criminal justice outcomes, and for whom?" *Journal of Policy Analysis and Management*, 38(3): 600–628.

Molloy, Raven, Christopher L. Smith, and Abigail Wozniak. 2011. "Internal Migration in the United States." *Journal of Economic Perspectives*, 25(3): 173–196.

Myers, Samuel L. 1980. "Why are Crimes Underreported? What is the Crime Rate? Does it "Really" Matter?" *Social Science Quarterly*, 61(1): 23–43.

Notowidigdo, Michael J. 2020. "The Incidence of Local Labor Demand Shocks." *Journal of Labor Economics*, 38(3): 687–725.

Nunn, Nathan, and Nancy Qian. 2014. "US Food Aid and Civil Conflict." *American Economic Review*, 104(6): 1630–1666.

Ong, David, and Jue Wang. 2015. "Income Attraction: An Online Dating Field Experiment." *Journal of Economic Behavior & Organization*, 111: 13–22.

Poterba, James M, Steven F Venti, and David A Wise. 2000. "Saver behavior and 401 (k) retirement wealth." *American Economic Review*, 90(2): 297–302.

Quadrini, Vincenzo. 2020. "The impact of industrialized countries? monetary policy on emerging economies." *IMF Economic Review*, 68(3): 550–583.

Ravenstein, Ernest George. 1885. "The Laws of Migration." *Journal of the Royal Statistical Society*, 48: 167–235.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2020. IPUMS USA: Version 10.0 [dataset]. Minneapolis, MN.

Saint-Paul, Gilles. 2015. "Genes, Legitimacy and Hypergamy: Another Look at the Economics of Marriage." *Journal of Demographic Economics*, 81(4): 331–377.

Shenhav, Na'ama. 2021. "Lowering Standards to Wed? Spouse Quality, Marriage, and Labor Market Responses to the Gender Wage Gap." *The Review of Economics and Statistics*, 103(2): 265–279.

Skogan, Wesley G. 1975. "Measurement problems in official and survey crime rates." *Journal of criminal justice*, 3(1): 17–31.

Skogan, Wesley G. 1977. "Dimensions of the dark figure of unreported crime." *Crime & Delinquency*, 23(1): 41–50.

Stolzenberg, Lisa, David Eitle, and Stewart J D'alessio. 2006. "Race, economic inequality, and violent crime." *Journal of Criminal Justice*, 34(3): 303–316.

Thaler, Richard H, and Shlomo Benartzi. 2004. "Save more tomorrow?: Using behavioral economics to increase employee saving." *Journal of political Economy*, 112(S1): S164–S187.

Tolbert, Charles M, and Molly Sizer. 1996. "U.S. Commuting Zones and Labor Market Areas: A 1990 Update." *Economic Research Service Staff Paper 9614*.

Weiss, Yoram, Junjian Yi, and Junsen Zhang. 2013. "Hypergamy, Cross-Boundary Marriages, and Family Behavior." *IZA Discussion Papers 7293*.

Witte, Ann Dryden. 1980. "Estimating the economic model of crime with individual data." *The quarterly journal of economics*, 94(1): 57–84.

Zhang, Hanzhe. 2021. "An Investment-and-Marriage Model with Differential Fecundity: On the College Gender Gap." *Journal of Political Economy*, 129(5): 1464–1486.