Male-biased Demand Shocks and Women's Labor Force Participation: Evidence from Large Oil Field Discoveries *

Stephan E. Maurer† and Andrei V. Potlogea‡

November 9, 2016

Job Market Paper
Latest version available here

Abstract

Do male-biased demand shocks affect women's labor force participation? To study this question, we examine large oil field discoveries in the US South from 1900-1940. We find that oil wealth has a zero net effect on female labor force participation due to two opposing channels. Oil discoveries increase demand for male labor in oil mining and manufacturing and consequentially raise male wages. This leads to an increased marriage rate of young women, which could have depressed female labor force participation. But at the same time, oil wealth also increases demand for women in services, which counterbalances the marriage effect and leaves women's overall labor force participation rate unchanged. Our findings demonstrate that when the nontradable sector is open to women, male-biased demand shocks in the tradable sector need not reduce female labor force participation.

JEL classification: R11, N50, J12, J16

Keywords: oil, structural transformation, female labor force participation, gender pay gap

*We thank Guy Michaels, Steve Pischke, Giacomo Ponzetto, Esteban Aucejo, Ghazala Azmat, Michel Azulai, Nitika Bagaria, Dan Black, Florian Blum, Zelda Brutti, Francesco Caselli, Vasco Carvalho, Alejandro Cuñat, Marta De Philippis, Jeremiah Dittmar, Richard Hornbeck, Kilian Huber, Yu-Hsiang Lei, Alan Manning, Andrea Matranga, Rachel Ngai, Raphaël Parchet, Frank Pisch, Roland Rathelot, Ferdinand Rauch, Pedro Souza, Jaume Ventura, Alwyn Young and seminar participants at the LSE, Oxford, the 2014 EDP Jamboree in Paris and the 15th Trans-Atlantic Doctoral Conference at LBS for their helpful comments and suggestions. To Guy Michaels, we are particularly indebted for also sharing his dataset on oil discoveries with us. All remaining errors are our own. Potlogea acknowledges financial support from the Spanish Ministry of Economy and Competitiveness. A previous version of this paper was circulated under the title *Fueling the Gender Gap? Oil and Women's Labor and Marriage Market Outcomes.*

†LSE and CEP, S.E.Maurer@lse.ac.uk

‡University of Edinburgh, Andrei.Potlogea@ed.ac.uk

1 Introduction

Across the globe and over time, differences in female labor force participation are substantial (Olivetti 2013). Across sectors, sex ratios vary considerably, even within countries. For example, in 2008, women accounted for only 9 percent of all the hours worked in construction in the United States, but for 76 percent of all the hours worked in health services (Ngai and Petrongolo 2016). Because of such differences in the occupational structure, sectoral shocks can be gender-biased and affect men and women very differently.

One such example is oil wealth. The discovery of oil or natural gas resources is a male-biased labor demand shock that leads to an increase in male wages. However, as argued by Ross (2008), oil wealth may not only be beneficial to men, but also detrimental to women's labor force participation: Dutch Disease effects can harm other tradable sectors that employ women and reduce demand for female labor. Moreover, oil discoveries increase male wages, which, in household models of labor supply, may reduce married women's labor supply (Ashenfelter and Heckman 1974, Becker 1981). Thus, both supply of and demand for female labor decrease, and as a result female employment decreases as well.

In this paper, we use the discoveries of large oil fields in the Southern United States between 1900 and 1940 as a series of large and exogenous male-biased demand shocks to study the implications of such shocks for female labor force participation and for the marriage market. We use data on the location and discovery of large oil fields, coupled with individual-level census data, to compare the evolution of female labor market outcomes in oil-rich counties relative to baseline counties. Using this difference-in-differences strategy, we find that oil

discoveries have a zero net effect on female labor force participation. While oil shocks are initially male-biased and lead to increased male wages and male employment in oil mining and manufacturing, the service sector also expands and employs more women. If anything, we even find a slight increase in labor force participation for single women. At the same time, increased male wages increase the supply of "marriageable" men in the county and thus increase young women's marriage rates. Since married women were less likely to work, this has a depressing effect on female labor market involvement.

Our study suggests that the effect of oil wealth on women is not necessarily negative, but depends on social norms and institutions that establish whether other sectors are open to women and what the socially acceptable role of married women in the labor market is. In a broader context, our paper shows how the effects of a gender-biased shock to the tradable sector can be offset by income effects that lead to growth in the nontradable sector. We formalize this insight in a model of structural transformation in a small open economy. If women are allowed to work in services, our model shows that oil discoveries give rise to two opposing labor demand effects for females: A demand reduction in the tradable sector, and a demand increase in the nontradable sector. The more productive the oil industry, the more likely it is that the latter effect will dominate.

Our paper belongs to a sizeable literature that has analyzed the implications of resource shocks and other demand shocks for local economies and local labor markets (Pratt 1980, Corden and Neary 1982, Carrington 1996, Sachs and Warner 1999, 2001, Black, McKinnish and Sanders 2005, Papyrakis and Gerlagh 2007, Marchand 2012, Furchtgott-Roth and Gray 2013). We confirm the find-

ings of a series of a recent papers that have found positive effects of resource wealth on local economies (e.g. Michaels 2011, Aragon and Rud 2013, Fetzer 2014, Allcott and Kenniston 2014). Like them, we find that oil discoveries lead to sustained and substantial population growth and to a shift of the economy from agriculture into oil mining, manufacturing, and services. However, the focus of this literature has typically been on the more general implications of resource wealth in terms of employment and welfare, and gender differences have not been analyzed in much detail.

At the same time, the vast changes in female labor force participation over the last one hundred years have sparked considerable interest among labor economists. Research has highlighted the importance of both supply-side and demand-side factors as well that of changing social norms. Reasons for supply shifts include medical advances (Goldin and Katz 2005, Albanesi and Olivetti 2016), labor-saving technological progress in the production of household goods (Greenwood, Seshadri and Yorukoglu 2005, Jones, Manuelli and McGrattan 2015), declining child care costs (Attanasio, Low and Sanchez-Marcos 2008) and the persistent effects of male wartime mobilization (Acemoglu, Autor and Lyle 2004). On the demand-side, the literature has documented the importance of structural transformation (Akbulut 2011, Voigtlaender and Voth 2013, Ngai and Petrongolo 2016) and trade integration (Do, Levchenko and Raddatz 2016, Gladdis and Pieters 2016) as shifters of gender-specific demands. Finally, Fernandez (2011, 2013) and Alesina, Giuliano and Nunn (2013) present evidence for the importance of social norms. Klasen and Pieters (2015) analyze the determinants of continued low female labor force participation in India, and find evidence of adverse factors operating through both demand-side and supply-side channels,

as well as through conservative social norms.

To the best of our knowledge, Michael Ross (2008, 2012) was the first to connect resource wealth to female labor force participation. Several papers have followed since then (Kang 2009, Norris 2010, Kotsadam and Tolonen 2016). However, the literature has so far either focused on cross-country comparisons, or on short-run mining booms. The former suffer from the potential problem of not being able to control satisfactorily for differences in institutions, culture and other unobservables at the country level. The latter are informative for short-run changes, but less so about gender-biased shocks that have medium- to long-run consequences for sectoral composition. However, such longer lasting changes in sectoral composition are likely to lead to sustained changes in demand for female labor and are therefore particularly relevant for the question we study. Our setting is able to address both these issues.

Firstly, we rely on within-country variation in oil wealth that is plausibly exogenous. This allows us to avoid many of the shortcomings of cross-country studies such as the difficulty to adequately control for differences in institutions and culture. Using information about the timing of oil field discoveries, we can employ a clean difference-in-differences research design, comparing the evolution of oil-rich counties before and after discoveries to that of non-oil counties over the same time period. Oil discoveries as a source of variation in labor demand are also not subject to endogeneity concerns that arise in connection with other measures such as contemporaneous oil production.

Secondly, the oil field discoveries we study are large events with long-lasting consequences: they lead to sustained growth and structural transformation away from agriculture and into oil mining, services and manufacturing. Thus, our

setting is arguably closer to the theoretical mechanisms originally described by Ross (2008). Closest to our paper in this respect is the study by Aragon, Rud and Toews (2016), which looks at the collapse of the UK coal industry as a demand shock that is biased against males.

Finally, the case of the Southern United States offers us the advantages of analyzing consistent, high-quality data from a region that at that time still relied heavily on agriculture. This is important because much of the discussion surrounding the link between resource abundance and gender issues has concentrated on developing, often heavily agrarian economies.

The rest of this paper is organized as follows: Section 2 discusses the channels through which oil discoveries could influence women's labor market outcomes and presents a simple model of sectoral reallocation in the wake of oil discoveries. The data and empirical strategy of our paper are then discussed in section 3. Section 4 presents the results and several robustness checks, section 5 concludes.

2 Theoretical Mechanisms

Before we proceed to the main body of our analysis, this section provides a brief discussion of the mechanisms through which oil wealth could affect the labor market outcomes of women. Existing work (see for example Ross 2008 and 2012), has identified both demand side and supply side channels that could lead to mineral wealth having an adverse impact on women's involvement in the labor market.

2.1 A Model of Structural Transformation and Labor Demand

On the labor demand side, oil discoveries are likely to lead to substantial structural transformation in the local economy that can change demand for male and female labor. This has already been noted by Ross (2008) who describes how oil wealth can lead to a "Dutch Disease" that reduces demand for female labor by crowding out tradable goods industries that employ women. In what follows, we present a simple model of structural transformation in the wake of oil discoveries. It will not only guide our empirical analysis, but also allow us to speak about the implications of our findings beyond our empirical setting.

2.1.1 Set-up

Consider a local economy (county) endowed with exogenous populations of men and women (denoted L_M and L_W respectively). This local economy can participate in the production of two tradable goods, a consumption good Y_c (we may think of this as an agricultural good), and oil (Y_o). It also produces a nontradable good called local services (Y_s).

The consumption good is produced by combining male labor, female labor, and sector specific capital (we can think of this as land) via the Cobb-Douglas technology:

$$Y_c = A_c(L^c_M)^{\alpha}(L^c_W)^{\beta}(N_c)^{1-\alpha-\beta} \tag{1}$$

where A_c denotes the county's productivity in the delivery of the consumption good, L^c_M and L^c_W denote male and female labor employed in the production of the consumption good, while N_c represents the quantity of capital specific to

the consumption good employed in production. We assume that the county has an exogenous endowment of such capital (land) given by $\bar{N}_c$.

Oil, on the other hand can be produced using only male labor via the linear technology:

$$Y_o = A_o L_M^o \tag{2}$$

where A_o reflects the county's productivity in the oil sector, while L_M^o denotes the number of male workers employed in oil production.

Services, on the other hand, are produced using female labor via the linear technology:

$$Y_s = A_s L_W^s \tag{3}$$

where A_s reflects the county's productivity in the services sector, while L_W^s denotes the number of women employed in services.

We consider counties to be small open economies, and as a result the prices of tradable goods are exogenously determined: p_c for the consumption good, p_o for oil. The price of (nontradable) services, on the other hand, is determined locally and is endogenous. We denote it as p_s. We assume all good and factor markets are perfectly competitive.

Preferences of all agents in the economy (men, women and capital owners) are given by the Cobb-Douglas utility function:

$$U(c, s) = c^{\gamma} s^{1-\gamma} \tag{4}$$

where c denotes the quantity consumed of the consumption good while s denotes the quantity of services consumed.

Aside from working in the tradable sector or the services sector, women have access to an outside option given by domestic activities. This outside option yields an exogenous level of utility $\bar{u}$. Men don't have an outside option, so they will always be working, either in the tradable sector or in oil. [1]

In our empirical exercise we analyze a historical time period featuring relatively low female labor force participation. To make the model relevant to this set-up we assume that the outside option provided by the domestic activity is sufficiently attractive such that some women don't work, at least before oil is discovered (i.e. when oil productivity is zero or very low). We thus impose the restriction

$$\bar{u} > \left\{ \frac{\beta \gamma^{\beta} A_c (L_M)^{\alpha} (\bar{N}_c)^{1-\alpha-\beta}}{\left[\frac{\beta L_W}{1-\gamma+\gamma\beta}\right]^{1-\beta}} \right\}^{\gamma} [(1-\gamma)A_s]^{1-\gamma} \tag{5}$$

With the set-up above, we model oil discoveries as an exogenous increase in the productivity of the oil sector. In what follows, we explore the effects of rising productivity in the oil sector on the labor market outcomes of women. In particular, we are interested in the impact of oil discoveries (modeled as increases in A_o) on women's labor force participation. The main prediction of the model is presented in Proposition 1 below. The proof of this result is presented in Appendix A.

2.1.2 Main Result

Proposition 1. *There exist three thresholds of productivity in the oil sector, A_1^*, A_2^*, A_3^*, with $A_1^* < A_2^* < A_3^*$ such that:*

[1] The reason for this assumption is twofold. Firstly, we are primarily interested in modeling female labor force participation. Secondly, the average male labor force participation rate in our sample is around 90%, making it also an empirically less relevant case to study in our setting.

1. *If $0 \leq A_o < A_1^*$ no oil is produced in the county in equilibrium. Along this range, any increase in A_o has no effect on women's labor force participation (or wages). Increases in A_o along this range also do not affect male wages.*

2. *if $A_1^* < A_o < A_2^*$ oil is produced in the county in equilibrium. Along this range, any increase in A_o leads to a reduction in women's labor force participation but keeps women's wages unchanged. Increases in A_o along this range lead to an increase in male wages and to the proportion of men working in oil.*

3. *if $A_2^* < A_o < A_3^*$ oil is produced in the county in equilibrium. Along this range, any increase in A_o leads to an increase in women's labor force participation but keeps women's wages unchanged. Increases in A_o along this range lead to an increase in male wages and to the proportion of men working in oil.*

4. *if $A_o > A_3^*$ oil is produced in the county in equilibrium. Along this range, women's domestic outside option is no longer binding, and all women work in equilibrium. As a result, further increases in A_o are no longer associated with increases in women's labor force participation (which is now 100%), but to increases in female wages. Increases in A_o along this range lead to an increase in male wages and to the proportion of men working in oil.*

Corollary 1. *In the absence of the service sector (i.e. if $\gamma = 1$), any increase in A_o beyond the threshold at which the production of oil is economical in the county is associated with declines in women's labor force participation.*

Proof: See Appendix A.

For small values of A_o (i.e. before oil discoveries), the production of oil is uneconomical in the county. Moreover, in this range of low oil productivities, small

increases in oil productivity have no effect, as oil remains too costly to produce in the county. As oil productivity crosses the threshold A_1^*, the county becomes productive enough in oil to begin production. This puts upward pressure on male wages, as the oil sector contributes to the demand for male labor. Rising male wages in turn affect the demand for female labor and implicitly women's labor force participation. The link between male wages and female labor demand operates through two channels, one via the tradable sector and one via the services sector.

Within the tradable sector, an increase in male wages has two opposite effects on the demand for female labor. First, it tends to crowd out the tradable sector, as the costs of one of its inputs go up. We call this a scale effect, which leads to the loss of female employment in this sector. Second, the increase in male wages tends to encourage the substitution of male labor with female labor, thus increasing demand for the latter. The strength of this substitution effect depends on how easy it is to replace male with female labor (i.e. the greater the substitutability of male and female labor, the stronger this effect is). In the case of our Cobb-Douglas technology it can be shown that the scale effect is guaranteed to dominate the substitution effect, such that rising male wages are associated with declines in labor demand coming from the tradable sector[2].

The increase in male wages caused by growth in oil related activities also raises the demand for local services. This income effect tends to increase female employment in the service sector, thus potentially counterbalancing the depressing effect of oil on demand for female labor in the tradable sector. The

[2]For a more general specification in which male and female labor are combined into a labor aggregate with an elasticity of substitution given by ϵ it can be shown that the scale effect dominates the substitution effect as long as $\epsilon < \frac{1}{\alpha+\beta}$.

strength of this effect crucially depends on the expenditure share (i.e. size) of the service sector. The larger the share of income spent on local services, the more likely it is that growth in this sector compensates for loss of female employment in the tradable sector.

For an initial range of oil productivities beyond A_1^* (and up to A_2^*), the overall negative effect of growth in the oil sector on tradable sector female labor demand dominates growing female employment in the services sector. As a result, increases in A_o lead to decreases in female labor force participation. Beyond A_2^*, the relationship between these two forces is reversed, with the service sector employment growth dominating female job losses in the tradable sector. As a result, beyond A_2^* increases in A_o are associated with increases in female labor force participation. Finally, beyond A_3^*, male wages become sufficiently high such that high demand in the services sector draws the entire female population into the labor force (the outside option of domestic activity no longer binds). For $A_o > A_3^*$ further increases in A_o are no longer reflected in higher female labor force participation (which is now 100%), but in higher female wages.

Given the analysis above, it should be clear that in the absence of the service sector, any increase in A_o beyond the threshold that makes oil production economical in the county is associated with declines in female labor force participation. This is because the negative effect of rising male wages on the tradable sector is still present, while the compensating force of increasing female labor demand due to growth in services is no longer operational.

Finally, care has to be taken when taking the predictions of this model to the data. While we have modeled oil discoveries as a continuous increase in A_o, the setting we study is perhaps better described by a discrete jump in A_o. In this

context, the predictions of the model regarding the impact of oil discoveries on female labor force participation are ambiguous. As the relationship between (the continuous) A_o and female labor force participation is U-shaped, a discrete jump in A_o may be associated with a decline, an increase or no change in female labor force participation, depending on which portion of the U-shape the local economy jumps to.

Our model offers several important insights. Firstly, if the local service sector is absent or if women are precluded from working in it, oil discoveries reduce demand for female labor in the tradable sector and therefore drive them out of the labor market. On the other hand, if the service sector is present and is open to women, oil discoveries continue to harm demand for female labor in the tradable sector, but increase demand for female labor in the nontradable sector. The more productive the oil sector is, the more the latter effect is likely to dominate, and female labor force participation may even increase.

2.2 Reducing Labor Supply

In addition to these demand side effects, oil can also depress women's incentives to participate in the workplace, leading to an inward shift of their labor supply curve. Resource mining and its associated activities tend to display heavily male labor forces (something that will also be apparent in our setting) and thus constitute a male-biased labor demand shock. Such a shock translates into substantially higher male wages, which constitutes an increase in non-labor income for the spouses of these males. Moreover, in a Becker (1981) model of

household labor supply, an increase in male wages would increase men's comparative advantage in market work. Both arguments give rise to a negative elasticity of married women's labor supply with respect to their husband's wages, which is also consistent with a sizeable empirical literature in labor economics (e.g. Ashenfelter and Heckman 1974). In addition to boosting the non-labor income of the wives of oil workers, oil could also increase government revenue and thus government transfers. (Ross, 2008). However, recent studies have not found that oil wealth translates into substantial increases in public goods provision (Caselli and Michaels 2013, Martinez 2016). The more important effect would thus be on married women whose husbands experience wage increases. For this group, we would expect a decrease in labor supply.

Taken together, our theoretical discussion indicates two effects of oil on female labor supply: As shown in the previous section, the labor demand effect is ambiguous and depends on the size of the oil discovery shock. In addition, there is a negative labor supply effect for women due to their husbands' wages increasing. The overall effect is therefore ambiguous, and we proceed to assess it empirically.

3 Data and Empirical Strategy

In order to identify major oil field discoveries, we draw on the dataset compiled by Guy Michaels (2011), which lists all counties in the Southern and Southwestern United States that are situated above an oil field of 100 million barrel or more before any oil was extracted. We will refer to these counties simply as oil-rich counties. In addition, the dataset also contains information about the earliest

major discovery in each county. We treat a county as having discovered its oil wealth after the first major discovery has taken place. Figures 1 to 5 give an overview over the geography of oil discoveries between 1900 and 1940. There are only relatively few oil discoveries in the early years, with most discoveries happening in the 1920s and 1930s. In terms of geographic scope, we follow Michaels (2011) and include all counties within 200 miles of the oil-rich counties in Texas, Louisiana, and Oklahoma. This gives us a sample of 774 counties in total.[3]

For our outcome variables, we use the individual-level US census data available from the Integrated Public Use Microdata Series (IPUMS) for 1900 to 1940. This includes three novel, recently made available full count (100%) samples for 1920, 1930, and 1940. For 1900 and 1910, we use the largest samples available- a 5% sample for 1900, and a 1.4% sample that oversamples minorities for 1910. We generally focus on the part of the population which is of working age, defined to be between 17 and 65.

Most of our variables of interest are available for our whole period of analysis. This includes labor force participation, the sector in which an individual works, sex, race, and marriage status. Employment status conditional on labor force participation was not recorded in the census of 1900 and 1920, but is available for the remaining three census decades. For earnings, we observe annual wage income in 1940 only, but we also have an "occupational earnings score" variable

[3]One potential issue is that especially in the Western part of the sample, population was growing and counties therefore often changed, were split up, or newly created. In order to address this, we compare the area of each county in each census decade to its area in 1940, and drop all observations from a given county-year cell if the absolute difference of the county's area in that year compared to 1940 exceeds 5%. In addition, we drop all counties from Oklahoma in 1900, as this state at this point was still largely unorganized and divided into the Oklahoma and Indian Territory.

that is available for 1900-1940. This variable assigns each individual the percentile rank of its occupation in terms of median earnings in 1950.

The staggered nature of oil discoveries across space and over time lends itself quite naturally to a difference-in-differences research design. The basic regression we run uses the IPUMS individual level data and is of the form

$$y_{ct} = \alpha_c + \tau_t + \beta DiscoveredOilField_{ct} + X_c'\gamma_t + u_{ct} \tag{6}$$

where y_{ct} denotes outcome y (e.g. the labor force participation rate) in county c and year t. $DiscoveredOilField$ is a dummy that equals 1 if county c is oil-rich and if at least one of its major oil fields has already been discovered, and 0 otherwise. τ and α are year- and county fixed effects. X is a vector of control variables that vary at the county level. In line with Michaels (2011), we control for several geographic features that might be spuriously correlated with oil wealth: Longitude, latitude, aridity, average annual rainfall, distance to the closest navigable river, and distance to the ocean. All of these variables vary at the county level only, but we allow for them to have time-varying effects.

Since our key variation lies at the county x year level, we aggregate the individual census data to this level.[4] We then run our regressions at this level, weighting by the respective cell sizes and clustering standard errors at the county level. This approach produces identical point estimates and standard errors as running all our regressions at the individual level (Angrist and Pischke 2009). In all regressions, we cluster standard errors at the county level. When calculating standard errors, we use Young (2016)'s procedure that calculates an effective

[4] When aggregating, we weight individual observations by the person weights provided by IPUMS in order to improve the representativity of our sample.

degrees of freedom correction for clustered and robust standard errors. These effective degrees of freedom are also used when performing hypothesis tests.

For wages, we only have data from 1940, requiring a cross-sectional specification with stronger identifiying assumptions:

$$y_c = \beta DiscoveredOilField_c + X_c'\gamma + s_s + u_c \tag{7}$$

In this case, $DiscoveredOilField_c$ codes whether county c has an oil field which was already discovered by 1940. X is the same vector of control variables as above, and s are state fixed effects. The more demanding identification assumption behind this specification is that counties without oil wealth and counties that had not yet discovered their oil wealth by 1940 are a valid counterfactual for those counties with already discovered oil fields in 1940. In Appendix B, we show several robustness checks, the results of which support this assumption.

The setting of our analysis is the American Southwest five to nine decades after the American Civil War. After the war, the US South remained a predominantly agricultural region that lagged considerably behind the rest of the country and caught up only very slowly (Wright 1986, Caselli and Coleman 2001). This is also borne out by the county-level summary statistics displayed in table 1. We show the labor force shares of 4 major sectors (agriculture, manufacturing, services, oil mining) for the beginning and end of our sample. As can be seen, agriculture is by far the most important sector in 1900, but loses importance over the next 40 years. Manufacturing and services, on the other hand, are slowly gaining ground. Over the whole time period, our area of observation is a very rural region, with an urban population share of only 15%. In addition, the sum-

mary statistics also show a very low level of female involvement in the labor force. Female labor force participation stood at 16.6% in 1900 and increased by 3 percentage points by 1940. This is no Southern peculiarity: Across the whole United States, female labor force participation rose from 20.6% in 1900 to 25.8% in 1940 (Goldin 1990, Chapter 2). However, there was considerable heterogeneity along two dimensions: Black women were much more likely to work than white women (Goldin 1977), and single women had much higher labor force participation than married ones (Goldin 1990).

All in all, our basic setting is thus a still heavily rural and agricultural economy that is slowly moving towards more "modern" sectors, and which over the whole period of analysis displays relatively low female labor force participation rates.

4 Results

4.1 Oil and gender labor market differences

In 1900, the economy of the Southern United States was thus still predominantly agricultural and nearly no major oil field had yet been discovered, as can be seen from figure 1. To set the stage for our main analysis relating resource shocks to gender outcomes, we first check whether oil discoveries indeed represent important developments for the discovering counties. Table 2 shows difference-in-differences results for the log of the working age population, female population share, and the share of the labor force employed in, respectively, oil mining, manufacturing, services, and agriculture. We do not find any evidence for a

resource curse, but rather detect substantial growth: Population increases by around 42%, and the labor force shifts from agriculture into oil mining and, to a lesser extent, manufacturing and services. These findings largely confirm the results of Michaels (2011) for the same region, but are based on weaker identifying assumptions. They show that oil discoveries constitute substantial shocks to a county's economy that lead to growth and structural transformation. Interestingly, the population growth seems gender-balanced, with the female population share staying unchanged. In tables 3 and 4, we then analyze whether the economy's structural transformation affects men and women differentially. The results show that this is clearly the case: Oil discoveries lead to a decrease in the share of the male labor force employed in agriculture by 8.8 percentage points, while increasing the male labor force share employed in oil mining and manufacturing by 6.5 and 1.2 percentage points, respectively. As there is virtually no oil mining prior to oil discoveries, this means that all of a sudden, 6.5% of the whole male labor force work in the oil extraction industry, a sizeable change to the local economy's structure. Women, on the other hand, are nearly not affected at all by the oil industry. However, they also experience a decrease in the importance of agriculture of 6.3 percentage points, whereas the service sector's labor force share grows by 5.2 percentage points. Thus, both genders leave agriculture in relative terms, but while men go increasingly into the oil and manufacturing industries, women flock to the service sector. In terms of our model, it seems that the service sector in our economy is clearly open towards women.

In this case, our model has an ambiguous prediction in terms of female labor supply depending on the strength of the oil shock. Table 5 shows difference-

in-differences estimates of the effect of oil wealth on labor force participation by gender. In columns 1 and 3, we show results for male and female labor force participation, while columns 2 and 4 show results for the employment rate conditional on being in the labor force, a measure which we do not have for 1900 and 1920. Interestingly, we do not find large effects on either gender: While the male labor force participation seems to increase at least slightly in oil-rich counties, we do not observe any changes in female labor force participation. Similarly, we fail to identify any effect of oil discoveries on either the male or female employment rates. The respective point estimates are close to zero across all specifications and statistically insignificant. Thus, in contrast to the cross-country results of Ross (2008), we do not find evidence for a negative relationship between oil wealth and female labor market participation. The most likely explanation for this is the growth of the service sector. As discussed in section 2, even though the initial demand shock of oil is male-biased (and columns 1 of tables 3 and 4 suggest it indeed is), oil only reduces women's labor market involvement if there is no other sector that benefits from oil and is open to women. In our case, as table 4 shows, the service sector is playing this role: As the economy grows, demand for services increases, and more women go to work in such jobs.[5]

So far, we have focused on quantities transacted in the labor market. To analyze the effect of oil discoveries on wages, we look at two outcome variables. The first is the most direct measure of wages, annual wage income. However, this

[5]In Appendix B, we show that nearly the whole growth in the importance of the service sector as employer of female labor is due to the personal service and entertainment sector. This suggests that income effects rather than demand linkages from other industries are the driver of this growth.

variable is only available for 1940, which requires a cross-sectional specification with stronger identifying assumptions. To at least partially remedy this, we also look at the occupational earnings score, available from IPUMS. The occupational earnings score is a percentile rank of occupations based on their 1950 median earnings. It thus does not capture within-occupation changes in wages, but it does capture movements of people into jobs that pay relatively more or less. The advantage of this variable is that it is available for the whole time period of 1900-1940 so that we can use our difference-in-differences specification. In panel A of table 6, we report the results. In columns 1 and 2 we report the averages of log annual wage earnings for men and women separately. Unlike our previous results on labor force participation and employment, we identify substantial effects of oil wealth on local wage rates that differ by gender. While male wages increase by 26 log points on average, female wages stay nearly unchanged. In columns 3 and 4, we convert annual wage income to percentile ranks of the 1940 wage distribution to facilitate comparison with the occupational score variable.[6] We see the same relative movement: Men move up 7.5 percentiles, wherewas women increase their earnings by at most one percentile. Columns 5 and 6 display cross-sectional estimates for 1940 for the occupational score variables. They show the same relative picture: Men gain 5 percentage ranks, while women in this specification even lose ground. In the panel difference-in-differences specification of columns 7 and 8, the gap between the estimates is a bit smaller, but still sizeable and in line with previous findings. Clearly, the structural transformation that counties experience in the wake of

[6]In creating percentile ranks, we follow the IPUMS procedure behind occupational earnings scores very closely: We first use the 1940 national full count to create a z-score of wage income and then transform this into a percentile rank.

oil discoveries is not gender-neutral: While men obtain high-paying jobs in oil mining and manufacturing, the service sector jobs for women tend to be relatively low-paid occupations in the personal service sector. The gender gap therefore increases. In addition, comparing columns 3 and 4 to 5-8, we see that a large part of the wage growth is due to occupational changes, but there is also evidence for within-occupational wage increases, as the wage percentile coefficients are larger than then occupational score ones. One concern with these findings is that the male wage growth could be driven by the high wages in oil mining and thus only accrue to the small group of oil miners. In panel B we therefore drop all individuals working in the oil extraction sector. As can be seen, while the wage growth decreases a bit, it stays economically and statistically significant: Outside of the oil industry, men gain 3 percentile ranks in the occupational earnings distribution, 6 percentiles in the wage distribution and 22.5 log points in wages. For women, on the other hand, the coefficients are nearly unchanged. Given the low number of women in the oil mining industry, this should not come as a surprise.

All in all, so far, our results display a much more nuanced effect of the initial male-biased demand shock on women: While men greatly benefit from new jobs in oil mining and manufacturing and can substantially increase their wages, female wages remain more or less unchanged. Moreover, oil does not appear to drive women out of the labor market, instead female labor force participation remains unchanged. The likely explanation is an increase in service sector jobs that absorbs any "excess" female labor. However, the service sector increase comes predominantly in personal service sector jobs that are not as well paid, such that the gender pay gap increases.

4.2 Heterogeneity along marriage status and race

One potential issue with our findings so far, in particular with the absence of a negative effect on female labor force participation, could be that most women were not part of the labor force anyways. As Goldin (1990, 2006) has shown, female labor force participation was still very low and only slowly rising during the first third of the twentieth century. This is also visible in our sample, where the female labor force participation rate over the whole time period is just below 24%. The vast majority of women thus was most likely already in a corner solution with no labor force participation, which could explain why we do not observe any changes in labor supply. While this argument has merit, it overlooks two important sources of heterogeneity in female labor market involvement: Marriage status and race. We therefore analyze whether we observe heterogeneous responses to oil discoveries along these dimensions, and whether these potential heterogeneous responses can explain the absence of an overall change in female labor market involvement.

Marriage status was a crucial determinant of labor force participation in the early 20th century United States. Until the 1920s, it was typical for women to exit the labor force upon marriage (Goldin 2006), and in many places, employers had explicit "marriage bars" that prevented the hiring or led to the firing of married women. Not surprisingly then, in 1900, across the whole country, only 5.6% of all married women were working, as opposed to 43.5% of all single women, and the former share had only increased to 13.8% by 1940. (Goldin

1990) Our failure to find a significant labor force participation result could thus be due to us pooling irresponsive married women with more responsive singles. If this is the case, we would expect to find a reaction for singles at least. Table 7 therefore shows results separately for single and married women. As can be seen, while we find a small negative point estimate for married women, if anything, our point estimate for single women is positive and amounts to 1.3 percentage points. While not very precisely estimated, this would mean that far from crowding singles out of the labor market, oil if anything *increased* their propensity to work. For the more responsive group of singles, our conclusion that oil does not drive them out of the labor market thus holds a fortiori.

However, in addition to changing female labor force participation directly, oil might also have other demographic effects. As we have shown, male wages increase substantially after oil discoveries. This could increase the local supply of "marriageable" men and thereby increase marriage rates (Buckley 2003, Jelnov 2016). Increased marriage rates in turn have a dampening effect on female labor force participation due to married women being less likely to work.

One challenge in assessing the plausibility of this channel is the problem that there is a large "stock" of already married women who are not affected by oil discoveries at all. We therefore need to focus on "marginal" women whose decision when and whether to marry can still be affected by oil wealth- these will typically be younger women, among which the marriage rate and thus the stock of married women is lower. In table 8, we show the effects of oil discoveries on a woman's probability of being married. Column 1 shows results for all women and illustrates our stock problem: Among all women, the marriage rate is close to 70% (and many of those not married are widowed already), and we do not

find an effect of oil on marriage rates. For women aged 30 or below, or 25 or below, where more women are in the process of making their marriage decision, we find that oil increases the probability of being married by a statistically significant 1.5 percentage points.

Overall, separating the analysis for single and married women supports our previous conclusion that oil does not reduce female labor force participation. If anything, single women even experience a slight increase in labor market involvement. However, this effect is counterbalanced by the fact that young women also are more likely to get married, and married women during this period rarely work. Social norms about the appropriate role of married women thus seem to keep women from benefiting more from oil than could have been possible.

In addition to marriage status, race was another important determinant of female labor force participation in our setting. Black women were considerably more likely to work. In 1920, the labor force participation of white females was 21.6%, while for nonwhite women, it stood at 43.1% (Goldin 1990). Reasons for this difference include economic factors such as black husbands' lower average incomes, but also cultural differences and the legacy of slavery. (Goldin 1977, Boustan and Collins 2013)

Our overall unchanged female labor force participation rate thus might mask racial heterogeneity. Black women with a generally greater labor force involvement should be more responsive to the shocks brought about by oil discoveries. In addition, they were particularly present in the low-skilled personal service sector jobs that experienced particular employment increases. (Goldin 1977) To assess whether black and white women react differentially to oil discoveries,

we split our sample by race. Table 9 shows the results for labor force participation, average earnings scores and the share of the labor force working in services and agriculture- in columns 1-4 for whites, in columns 5-8 for blacks. The results show that black women indeed react much more to oil discoveries: The importance of the agricultural sector as employer decreases by nearly 9 percentage points for them, while the services experience a corresponding increase. This movement from agriculture to services also appears financially beneficial, as black women gain more than a percentile rank in the occupational earnings score. For white women, on the other hand, all these responses are qualitatively of the same sign, but much more muted. The likely explanation for this heterogeneity is that white women start from considerably better paying jobs- they are on average in the 32nd percentile of the median occupational earnings distribution, whereas blacks are in the 7th, and few white women work in agriculture. For whites, who already have relatively well-paid service sector jobs, the new personal service jobs are thus not very attractive. Black women, on the other hand, can considerably improve over their existing agricultural jobs. Finally, however, overall labor force participation stays unchanged for both races.

All in all, analyzing the effect of oil along the dimensions of race and marriage status has uncovered several important additional results. First off, oil discoveries do increase the earnings of women, but only of black women who on average had the worst-paying jobs and for whom the new personal service sector jobs are an improvement. White women, the majority of which is already in the service sector, do not gain and also do not react a lot in general. Secondly, if anything, single women even experience an increase in labor market opportunities, but this effect is counterbalanced by an increased propensity to marry and a so-

cial environment where married women are strongly discouraged from working. Again, as in the case of the service sector's openness, this highlights the role of social norms. Had the labor market been more open to married women, we might even have seen a slight increase in female labor force participation. Instead, women seem to benefit less than they could have. Finally, our general results seem robust to splitting the sample of women both by race and marriage status. In Appendix B, we perform further robustness checks and show that our results are also robust to using a leads and lags specification instead of a simple dummy for oil discoveries, to dropping all counties without large oil fields, and to excluding all counties who do not have a discovered oil field themselves, but are adjacent to one.

4.3 Migration

As we have shown in table 2, counties that discover oil experience great population growth. This immediately raises the question of migration. Migration is not a problem for the identification strategy of our paper, but it poses potential challenges for the interpretation of our results: Do the changes and potential benefits of oil accrue only to migrants, or does the incumbent population also share the benefits? The ideal test of this hypothesis would be to split the sample by migrant status. Unfortunately, however, we do not observe migrant status. What we do know is each individual's state of birth. We can thus at least identify a subsample of migrants, namely those who migrated across state borders. These are also the migrants who on average traveled greater distances. If there is a concern of selective migration and spatial sorting, this would therefore be

the group where we would suspect these factors to operate in the strongest way. As one admittedly imperfect way of addressing the questions posed by migration, we therefore split the sample by birth- one group consists of people that were born in a different state and thus clearly migrated to their new residence at some point in their life (albeit not necessarily after oil discoveries). The second group consists of people that live in their state of birth. This group is made up of short-distance migrants and individuals who have not moved at all. In tables 10 and 11, we show our key results separately for these two groups. Table 10 shows results for the overall sex ratio and for the male-related outcomes: labor force participation rate, employment rate, average earnings scores and the sectoral allocation of the labor force. For men, we can indeed see some evidence for selective migration: For men born out of state, the occupational earnings score results and the sectoral reallocation from agriculture to oil and manufacturing are considerably stronger than for men born in their state of residence. However, neither group displays a huge change in labor force participation rate. Table 11 shows the results for our female-related outcome: Labor force participation and employment rates, marriage rates of women aged 25 or below, average occupational earnings scores and sectoral composition of the labor force. Here, the case for selective migration seems considerably weaker: The marriage rate increase is driven almost entirely by women born in state, and the sectoral reallocation from agriculture to services is very similar for both groups. In addition, again no group displays a change in labor force participation rates or earnings scores. Overall, for men, selective migration is clearly a concern and at least part of our findings are due to migrants being systematically different from the existing population. However, for women, this case seems much weaker.

5 Conclusion

In this paper, we have analyzed the implications of a male-biased labor demand shock in the tradable sector on labor market outcomes by gender. Our particular focus has been on testing the hypothesis that oil discoveries adversely affect female labor market participation.

We have provided a theoretical framework in which we model the oil sector as male-dominated and oil discoveries as male-biased labor demand shocks. In our model, oil discoveries lead to two opposing changes in female labor demand: Demand for female labor in the tradable sector is reduced, whereas demand for women in the nontradable sector increases. If the service sector does not exist or is not open to women, female labor force participation unambiguously declines. On the other hand, if the service sector is open to women, the nontradable sector might have enough pull to actually increase female labor force participation.

We then use oil discoveries in the US Southwest between 1900 and 1940 to empirically estimate the effect of oil wealth on gender labor market differences. Contrary to previous studies (e.g. Ross 2008), we do not find a negative relationship between oil wealth and female labor force participation, which is instead found to be unchanged. Our explanation for this is a process of sectoral reallocation by gender that is consistent with our model: The initial oil shock indeed brings about male-biased employment increases in oil mining and manufacturing with associated wage gains for men. However, the service sector also expands and offers ample employment opportunities for women, who therefore do not leave the labor force. If anything, we even find that the labor market

prospects of single women slightly improve. However, at the same time, there is an increased supply of marriageable men in the county, and the marriage rate of young women increases. In our period of analysis, married women were strongly discouraged from market work, so that this increased marriage rate depresses any potential gains women could have made from oil discoveries.

Our findings point to the importance of social norms and institutions in transmitting gender-biased demand shocks. A first such norm is the openness of the service sector to women. In our setting, the service sector is open to women and expands considerably. Our model indicates that if the latter was not the case, we would indeed expect to find a negative employment effect for women. This could also explain the negative cross-country result in Ross (2008), which, as noted by himself, seems to be absent in countries where women are common in the service sector.

A second relevant social norm concerns the "appropriate" role of married women. In our setting, married women are socially discouraged (and sometimes outright barred) from working, and as a consequence, they potentially benefit less from oil discoveries than they could have under different social norms.

The broader implication of our paper is that an initial gender-biased shock to a tradable sector does not necessarily lead to aggregate gender-biased outcomes. Through income effects (or potentially backward or forward linkages), the initial gender-biased implications can easily be attenuated or even fully netted out.

References

Acemoglu, Daron, David H. Autor and David Lyle. 2004. Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury. *Journal of Political Economy* 112: 497-551

Akbulut, Rahsan. 2011. Sectoral changes and the increase in women's labor force participation. *Macroeconomic Dynamics* 15:240-64.

Albanesi, Stefania, and Claudia Olivetti. 2016. Gender Roles and Medical Progress. *Journal of Political Economy* 124: 650-695

Alesina, Alberto F., Paola Giuliano, and Nathan Nunn. 2013. On the origins of gender roles: Women and the plough. *Quarterly Journal of Economics* 128: 469-530.

Allcott, Hunt and Daniel Keniston. 2014. Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resourc Booms in Modern America. NBER Working Paper 20508

Angrist, Joshua and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics*. Princeton and Oxford: Princeton University Press

Aragon, Fernando M., and Juan P. Rud. 2013. Natural resources and local communities: Evidence from a peruvian gold mine. *American Economic Journal: Economic Policy* 5: 1-25.

Aragon, Fernado M., Juan P. Rud and Gerhard Toews. 2016. Resource shocks, employment, and gender: evidence from the collapse of the UK coal industry. Working Paper.

Ashenfelter, Orley and James Heckman. 1974. The Estimation of Income and Substitution Effects in a Model of Family Labor Supply. *Econometrica* 42:73-85

Attanasio, Orazio, Hamish Low, and Virginia Sanchez-Marcos. 2008. Explaining changes in female labor supply in a life-cycle model. *American Economic Review* 98: 1517-42.

Becker, Gary S. 1981. *A Treatise on the Family* Cambridge and London: Harvard University Press

Black, Dan, Terra McKinnish, and Seth Sanders. 2005. The economic impact of the coal boom and bust. *The Economic Journal* 115: 449-76.

Boustan, Leah Platt and William J. Collins. 2013. The Origin and Persistence of Black-White Differences in Women's Labor Force Participation In: Leah Platt Boustan, Carola Frydman, Robert A. Margo (eds). *Human Capital in History: The American Record*. Chicago: University of Chicago Press

Buckley, Patrick D. 2003. Essays in applied microeconomics: Signaling, marriage hazard rates, and ability sorting. PhD diss., Department of Economics, Massachusetts Institute of Technology.

Carrington, William J. 1996. The Alaskan Labor Market during the Pipeline Era. *Journal of Political Economy* 104: 186-218

Caselli, Francesco, and John W. Coleman. 2001. The U.S. structural transformation and regional convergence: A reinterpretation. *Journal of Political Economy* 109: 584-616.

Caselli, Francesco, and Guy Michaels. 2013. Do oil windfalls improve living standards? Evidence from Brazil. *American Economic Journal: Applied Economics* 5: 208-38

Corden, Max W., and Peter J. Neary. 1982. Booming sector and de-industrialisation in a small open economy.*The Economic Journal* 92: 825-48.

Do, Quy-Toan, Andrei A. Levchenko, and Claudio Raddatz. 2016. Comparative advantage, international trade, and fertility.*Journal of Development Economics* 119: 48-66

Fernandez, Raquel. 2011. Does culture matter?. In *Handbook of Social Economics*, vol. 1A, ed. Jess Benhabib, Matthew O. Jackson, and Alberto Bisin. Amsterdam: North Holland.

Fernandez, Raquel. 2013. Cultural change as learning: The evolution of female labor force participation over a century. *American Economic Review* 103: 472-500.

Fetzer, Thiemo R. 2015. Fracking growth. CEP Discussion Paper No. 1278, London School of Economics.

Furchtgott-Roth, Diana, and Andrew Gray. 2013. The economic effects of hydrofracturing on local economies: A comparison of New York and Pennsylvania. Growth and Prosperity Report no.1, Manhattan Institute for Policy Research, New York.

Gaddis, Isis, and Janneke Pieters. 2016. The Gendered Labor Market Impacts of Trade Liberalization: Evidence from Brazil. Forthcoming, *Journal of Human Resources*

Goldin, Claudia. 1977. Female Labor Force Participation: The Origin of Black and White Differences, 1870 and 1880. *Journal of Economic History* 37: 87-108

Goldin, Claudia. 1990. *Understanding the gender wage gap: An economic history of American women*. New York: Oxford University Press.

Goldin, Claudia. 2006. The quiet revolution that transformed women's employment, education, and family. *American Economic Review, Papers and Proceedings* 96: 1-21.

Goldin, Claudian and Lawrence F. Katz. 2002. The power of the pill: Oral contraceptives and women's career and marriage decisions. *Journal of Political Economy* 100: 730-70.

Greenwood, Jeremy, Ananth Seshadri, and Mehmet Yorukoglu. 2005. Engines of liberation. *Review of Economic Studies* 72: 109-33.

Haines, Michael R., and Inter-university Consortium for Political and Social Research. 2010. Historical, demographic, economic, and social data: The United States, 1790-2002. ICPSR02896-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. http://doi.org/10.3886/ICPSR02896.v3

Humphreys, Macartan, Jeffrey D. Sachs, and Joseph E. Stiglitz. eds. 2007. *Escaping the resource curse*. New York, NY: Columbia University Press.

Jelnov, Pavel. 2016. The Marriage Age U-shape: What Can we Learn from Energy Crises. Working Paper

Jones, Larry E., Rodolfo E. Manuelli, and Ellen R. McGrattan. 2015. Why are married women working so much? *Journal of Demographic Economics* 81: 75-114

Kang, Alice. 2009. Studying oil, Islam, and women as if political institutions mattered. *Politics & Gender* 5: 560-68

Klasen, Stephan, and Janneke Pieters. 2015. What explains the Stagnation of female labor force participation in urban India? *World Bank Economic Review* 29: 449-478

Kotsadam, Andreas, and Anja Tolonen. 2016. African mining, gender, and local employment. Forthcoming, *World Development*

Marchand, Joseph. 2012. Local labor market impacts of energy boom-bust-boom in Western Canada. *Journal of Urban Economics* 71: 165-74.

Martinez, Luis R. 2016. Sources of Revenue and Government Performance: Evidence from Colombia. Working Paper.

Michaels, Guy. 2011. The long term consequences of resource-based specialisation. *The Economic Journal* 121: 31-57.

Minnesota Population Center. 2011. *National historical geographic information system: Version 2.0*. Minneapolis, MN: University of Minnesota.

Ngai, Rachel L., and Barbara Petrongolo. 2014. Gender Gaps and the Rise of the Service Economy. Discussion Paper no. 8134, IZA, Bonn.

Norris, Pippa. 2010. Perhaps petroleum perpetuates patriarchy? A response and critique to Ross. Unpublished manuscript. Harvard Kennedy School of Government.

Olivetti, Claudia. 2013. "The Female Labor Force and Long-run Development: The American Experience in Comparative Perspective", Working Paper no. 19131, National Bureau of Economic Research, Cambridge, MA.

Penn Energy Research 2000. Oil and Gas Journal Data Book. Tulsa, OK: Pennwell Publishing Co.

Penn Energy Research 2006. Oil and Gas Journal Data Book. Tulsa, OK: Pennwell Publishing Co.

Papyrakis, Elissaios, and Reyer Gerlagh. 2007. Resource abundance and economic growth in the United States. *European Economic Review* 51: 1011-39.

Pratt, Joseph A. 1980. *The growth of a refining region*. Greenwich, Connecticut: Jai Press.

Ross, Michael L. 2008. Oil, Islam and women. *American Political Science Review* 102: 107-23.

Ross, Michael L. 2012. *The oil curse: How petroleum wealth shapes the development of nations*. Princeton, NJ: Princeton University Press.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder and Matthew Sobek. 2010. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota.

Sachs, Jeffrey D. and Andrew M. Warner. 1999. The Big Rush, Natural Resource Booms And Growth. *Journal of Development Economics* 59: 43-76

Sachs, Jeffrey D. and Andrew M. Warner. 2001. The curse of natural resources. *European Economic Review* 45: 827-38.

Voigtlaender, Nico, and Joachim Voth. 2013. How the West 'invented'fertility restriction. *American Economic Review* 103: 2227-64

Wright, Gavin. 1986. *Old South, new South*. New York:Basic Books.

Young, Alwyn. 2016. Improved, Nearly Exact, Statistical Inference with Robust and Clustered Covariance Matrices using Effective Degrees of Freedom Corrections. Working Paper

Tables and Figures

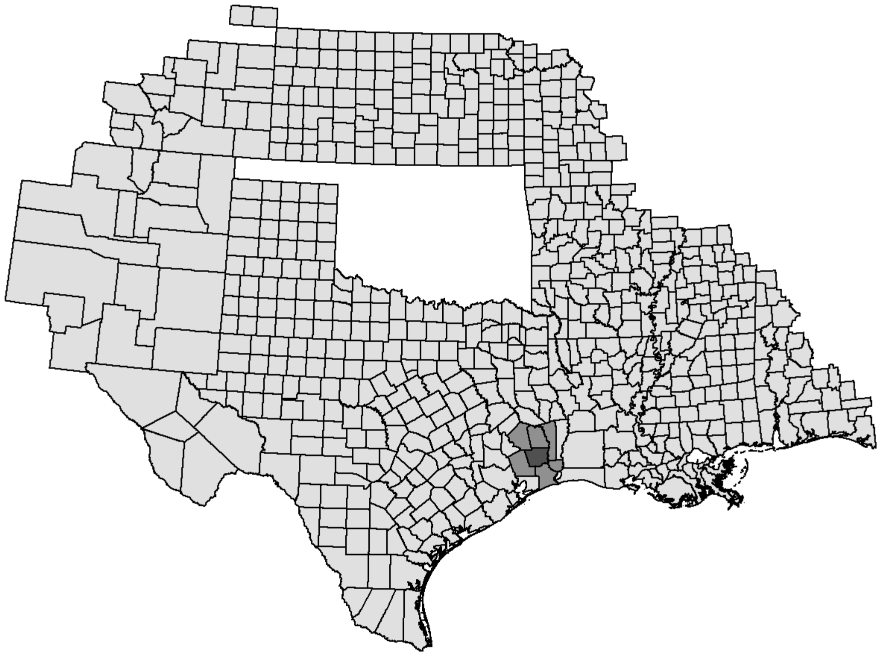

Figure 1: Map of Oil Discoveries 1900

Note: Oil abundant counties (dark grey), Neighbors of oil abundant counties (grey), Other counties (light grey)

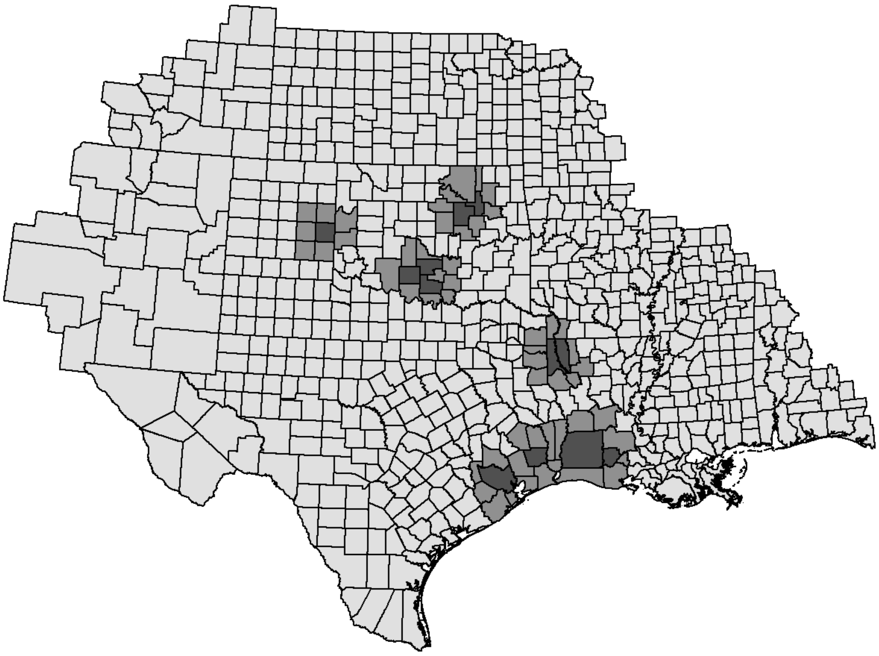

Figure 2: Map of Oil Discoveries 1910

Note: Oil abundant counties (dark grey), Neighbors of oil abundant counties (grey), Other counties (light grey)

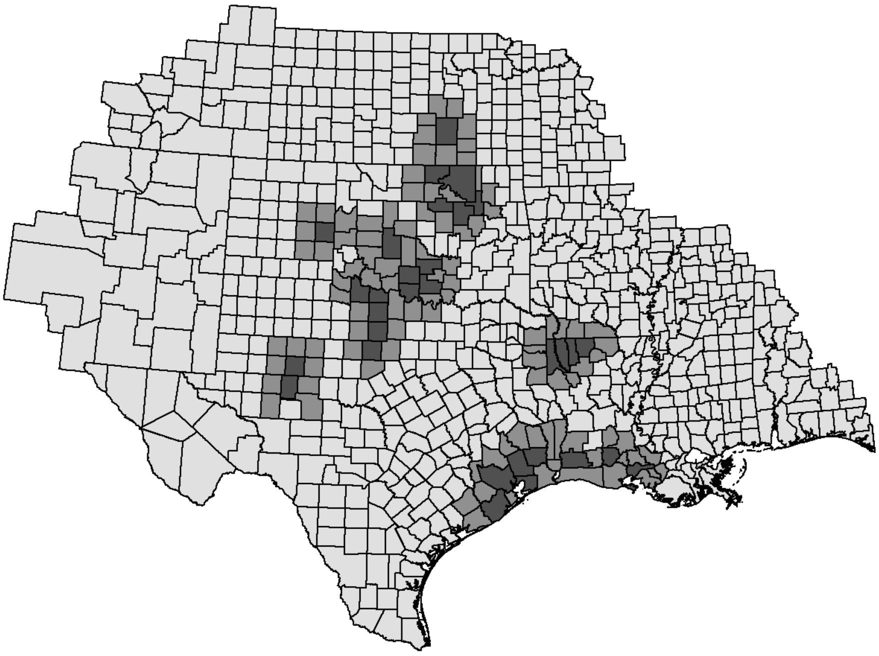

Figure 3: Map of Oil Discoveries 1920

Note: Oil abundant counties (dark grey), Neighbors of oil abundant counties (grey), Other counties (light grey)

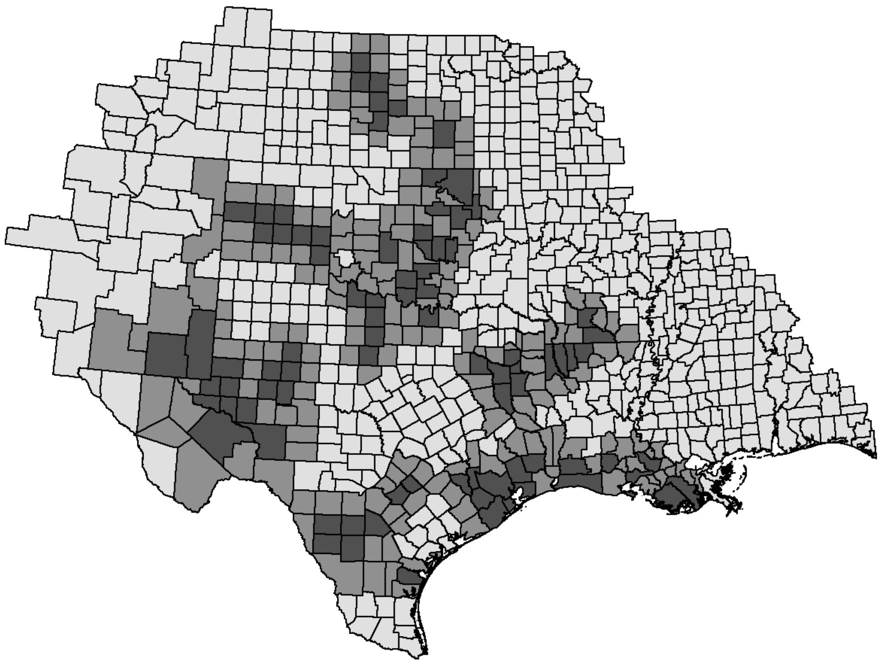

Figure 4: Map of Oil Discoveries 1930

Note: Oil abundant counties (dark grey), Neighbors of oil abundant counties (grey), Other counties (light grey)

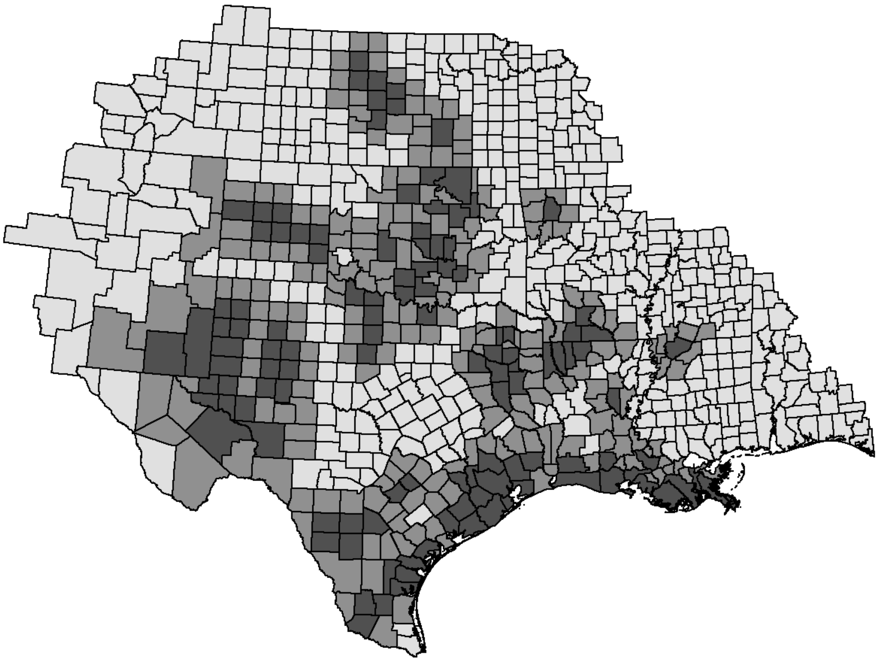

Figure 5: Map of Oil Discoveries 1940

Note: Oil abundant counties (dark grey), Neighbors of oil abundant counties (grey), Other counties (light grey)

	1900			1940		
	Mean	Standard deviation	Obs.	Mean	Standard deviation	Obs.
Agriculture LF Share	0.716	0.199	588	0.481	0.184	758
Manufacturing LF Share	0.038	0.054	588	0.062	0.057	758
Services LF Share	0.102	0.070	588	0.152	0.057	758
Oil Mining LF Share	0.000	0.001	588	0.016	0.045	758
Female LF Participation Rate	0.166	0.134	587	0.198	0.070	774
Male LF Participation Rate	0.934	0.045	588	0.866	0.141	774
Black population share	0.206	0.256	588	0.152	0.199	774
Urban population share	0.080	0.172	588	0.214	0.232	774
Marriage share, all women	0.692	0.092	587	0.727	0.039	774
Marriage share, women aged 16 to 25	0.495	0.150	581	0.516	0.071	774

Table 1: Summary statistics

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Ln(population)	Female pop share	Share of the LF employed in			
			Oil	Manuf.	Agriculture	Services
Discovered Oil Field	0.351***	-0.003	0.052***	0.008**	-0.083***	0.013**
	(0.068)	(0.003)	(0.007)	(0.004)	(0.017)	(0.006)
Observations	3,594	3594	3,577	3,577	3,577	3,577
Mean Dep Var		0.492	0.010	0.078	0.494	0.167

Number of clusters 774. Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** p<0.01, ** p<0.05, * p<0.1

Table 2: Broad economic effects of oil discoveries

VARIABLES	(1)	(2)	(3)	(4)
	Share of the Labor Force employed in			
	Oil Mining	Manufacturing	Agriculture	Services
Discovered Oil Field	0.065***	0.012***	-0.088***	0.004
	(0.008)	(0.004)	(0.016)	(0.004)
Mean Dep Var	0.012	0.087	0.536	0.087

Number of observations 3,577, number of clusters 774. Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** p<0.01, ** p<0.05, * p<0.1

Table 3: Sectoral shifts by gender: Men

VARIABLES	(1)	(2)	(3)	(4)
	Share of the Labor Force employed in			
	Oil Mining	Manufacturing	Agriculture	Services
Discovered Oil Field	0.001***	-0.009*	-0.063**	0.052*
	(0.000)	(0.005)	(0.030)	(0.028)
Mean Dep Var	0.001	0.045	0.329	0.479

Number of observations 3,475, number of clusters 774. Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** p<0.01, ** p<0.05, * p<0.1

Table 4: Sectoral shifts by gender: Women

VARIABLES	(1) LF Part. Rate Female	(2) Employment Rate Female	(3) LF Part. Rate Male	(4) Employment Rate Male
Discovered Oil Field	-0.003	-0.004	0.008	-0.005
	(0.010)	(0.008)	(0.007)	(0.004)
Observations	3,590	2,171	3,594	2,231
Mean Dep Var	0.246	0.938	0.907	0.941
Years	1900-1940	1910, 30, 40	1900-1940	1910, 30, 40

Number of clusters 774. Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** p<0.01, ** p<0.05, * p<0.1

Table 5: The effect of oil on labor force participation rates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Average ln (wage income)		Average wage inc. percentile rank		Average occupational earnings score			
Panel A	All sectors							
Discovered Oil Field	0.263***	0.047	7.427***	1.045	5.188**	-1.392*	4.246***	0.531
	(0.080)	(0.071)	(2.289)	(1.383)	(2.119)	(0.734)	(0.918)	(0.607)
Mean Dep Var	6.268	5.730	41.159	36.199	38.018	26.979	33.019	21.197
Panel B	Excluding individuals working in the oil mining sector							
Discovered Oil Field	0.224**	0.044	6.179**	0.978	4.174*	-1.447*	2.958***	0.497
	(0.087)	(0.071)	(2.452)	(1.384)	(2.227)	(0.738)	(0.771)	(0.607)
Mean Dep Var	6.251	5.728	40.729	36.162	37.632	26.956	32.683	21.182
Sample	Male LF	Female LF	Male LF	Female LF	Male LF	Female LF	Male LF	Female LF
Observations	758	758	758	758	758	758	3,577	3,478
Clusters	758	758	758	758	758	758	774	774
Year	1940	1940	1940	1940	1940	1940	1900-1940	1900-1940

Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6: Oil discoveries and wages/earnings

	(1)	(2)
VARIABLES	Female LF Participation Rate	
Discovered Oil Field	0.013	-0.008
	(0.013)	(0.009)
Sample	Single Women	Married women
Observations	3,543	3,590
Mean Dep Var	0.483	0.139

Number of clusters 774. Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** p<0.01, ** p<0.05, * p<0.1

Table 7: The effects of oil on female LFP by marriage status

	(1)	(2)	(3)
VARIABLES	Share of women married		
Discovered Oil Field	0.002	0.015**	0.015*
	(0.004)	(0.007)	(0.008)
Sample	All Women	Women ≤ 30	Women ≤ 25
Observations	3,590	3,582	3,573
Mean Dep Var	0.690	0.594	0.501

Number of clusters 774. Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** p<0.01, ** p<0.05, * p<0.1

Table 8: Oil discoveries and marriage rates

VARIABLES	(1) LF Part. Rate	(2) Share of Female LF in services	(3) Share of Female LF in agriculture	(4) Average occ. earnings score	(5) LF Part. Rate	(6) Share of Female LF in services	(7) Share of Female LF in agriculture	(8) Average. occ. earnings score
Disc Oil Field	-0.000	0.019	-0.038	0.328	0.004	0.085**	-0.088**	1.217***
	(0.007)	(0.019)	(0.026)	(0.884)	(0.021)	(0.038)	(0.036)	(0.439)
Sample	White	White	White	White	Black	Black	Black	Black
Observations	3,590	3,452	3,452	3,455	2,954	2,686	2,686	2,709
Clusters	774	774	774	774	746	711	711	718
Mean Dep Var	0.184	0.483	0.188	31.899	0.459	0.475	0.495	7.318

Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 9: The effects of oil discoveries on women by race

VARIABLES	(1) Female pop share	(2) LF Part Rate	(3) Employment Rate	(4) Average occ. earnings score
Panel A	People born in state of residence			
Disc Oil Field	-0.000	0.010	-0.005	3.191***
	(0.002)	(0.008)	(0.004)	(0.642)
Observations	3,572	3,557	2,206	3,539
Panel B	People born outside of state of residence			
Disc Oil Field	-0.005	0.005	-0.007	5.913***
	(0.004)	(0.008)	(0.007)	(1.725)
Observations	3,586	3,583	2,221	3,566
Sample	All	Men	Male LF	Male LF

VARIABLES	(5) Share of Male LF employed in Oil Mining	(6) Manufacturing	(7) Services	(8) Agriculture
Panel A	People born in state of residence			
Disc Oil Field	0.038***	0.007	0.004	-0.061***
	(0.005)	(0.005)	(0.003)	(0.013)
Observations	3,538	3,538	3,538	3,538
Panel B	People born outside of state of residence			
Disc Oil Field	0.105***	0.015**	0.005	-0.126***
	(0.014)	(0.006)	(0.006)	(0.028)
Observations	3,566	3,566	3,566	3,566

Number of clusters 774. Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** p<0.01, ** p<0.05, * p<0.1

Table 10: Comparing state natives and state migrants, part 1

VARIABLES	(1) LF Part Rate	(2) Share Married	(3) Employment Rate	(4) Average occ. earnings score
Panel A	People born in state of residence			
Disc Oil Field	-0.006	0.015*	-0.007	0.654
	(0.012)	(0.008)	(0.007)	(0.622)
Observations	3,557	3,533	2,114	3,373
Panel B	People born outside of state of residence			
Disc Oil Field	0.002	0.009	0.001	0.268
	(0.090)	(0.016)	(0.016)	(1.027)
Observations	3,573	3,456	2,094	3,374
Sample	All Women	Women ≤ 25	Female LF	Female LF

VARIABLES	(5) Share of Female LF employed in Oil Mining	(6) Manufacturing	(7) Services	(8) Agriculture
Panel A	People born in state of residence			
Disc Oil Field	0.001***	-0.004	0.053	-0.066*
	(0.000)	(0.005)	(0.032)	(0.032)
Observations	3,370	3,370	3,370	3,370
Panel B	People born outside of state of residence			
Disc Oil Field	0.002***	-0.022***	0.060*	-0.068*
	(0.001)	(0.006)	(0.029)	(0.036)
Observations	3,365	3,365	3,365	3,365

Number of clusters 774. Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** p<0.01, ** p<0.05, * p<0.1

Table 11: Comparing state natives and state migrants, part 2

A Appendix A: Theoretical results

An equilibrium of the economy described in section 2 is characterized by the system of equations (8) to (17):

$$L_M^o + L_M^c = L_M \tag{8}$$

$$L_W^c + L_W^s = L_W^{LF} \leq L_W \tag{9}$$

$$N_c = \bar{N}_c \tag{10}$$

$$\frac{\beta}{\alpha}\frac{L_M^c}{L_W^c} = \frac{w_W}{w_M} \tag{11}$$

$$\frac{\beta N_c}{(1-\alpha-\beta)L_W^c} = \frac{w_W}{r_c} \tag{12}$$

$$p_c = \frac{1}{A_c}\left(\frac{w_M}{\alpha}\right)^{\alpha}\left(\frac{w_W}{\beta}\right)^{\beta}\left(\frac{r_c}{1-\alpha-\beta}\right)^{1-\alpha-\beta} \tag{13}$$

$$p_o \leq \frac{w_M}{A_o} \tag{14}$$

$$p_s = \frac{w_W}{A_s} \tag{15}$$

$$(1-\gamma)\left[w_W L_W^{LF} + w_M L_M + r_c N_c\right] = w_W L_W^s \tag{16}$$

$$\gamma^{\gamma}(1-\gamma)^{1-\gamma}\frac{w_W}{p_c^{\gamma}p_s^{1-\gamma}} \geq \bar{u} \tag{17}$$

where beyond previously established notation w_i are the wages of workers of gender $i \in \{W, M\}$, r_c is the rental rate of specific capital while L_W^{LF} denotes the number of women in the labor force (which due to the presence of the outside option can be fewer than the total population of women).

Equations (8) to (10) are market clearing conditions in the various factor markets. Equations (11) and (12) are optimality conditions from the maximization problem of consumption good producers. Equations (13) to (15) are zero profit

conditions for producers of the consumption good, oil and services respectively. Equation (16) is an optimality condition from the consumer's utility maximization problem. Equation (17) represents the participation condition of women into the labor force.

A.1 Low Oil Productivity

It can be shown that for A_o small enough such that (14) does not bind, the condition in (5) guarantees that (17) binds[7]. Therefore we begin our analysis with very small A_o where we know that (14) does not bind and (17) binds. In such an equilibrium we have that:

$$L_M^{c*} = L_M \tag{18}$$

$$L_M^{o*} = 0 \tag{19}$$

$$w_W^* = \frac{p_c \bar{u}^{\frac{1}{\gamma}}}{\gamma[(1-\gamma)A_s]^{\frac{1-\gamma}{\gamma}}} \tag{20}$$

where (20) is obtained by plugging (15) in (17) and rearranging. With these results in place the group of equations (11) to (13) represent a system of three equations with three remaining unknowns. We first divide (11) by (12) and rearrange to obtain:

$$r_c^* = \frac{(1-\alpha-\beta)L_M}{\alpha \bar{N}_c} w_M \tag{21}$$

[7]This can be shown by contradiction. Assume there is an equilibrium where no men work in oil, $\bar{u}$ obeys (5) and (17) does not bind. This leads to a contradiction.

We plug (20) and (21) into (13) and then solve for w_M yielding:

$$w_M^* = \alpha(A_c)^{\frac{1}{1-\beta}} p_c \left[\frac{\beta\gamma[(1-\gamma)A_s]^{\frac{1-\gamma}{\gamma}}}{\bar{u}^{\frac{1}{\gamma}}} \right]^{\frac{\beta}{1-\beta}} \left(\frac{\bar{N}_c}{L_M} \right)^{\frac{1-\alpha-\beta}{1-\beta}} \tag{22}$$

This configuration is only an equilibrium if we can confirm that (14) does not bind. Plugging (22) into (14) yields us with the condition

$$A_o < \underbrace{\alpha(A_c)^{\frac{1}{1-\beta}} \frac{p_c}{p_o} \left[\frac{\beta\gamma[(1-\gamma)A_s]^{\frac{1-\gamma}{\gamma}}}{\bar{u}^{\frac{1}{\gamma}}} \right]^{\frac{\beta}{1-\beta}} \left(\frac{\bar{N}_c}{L_M} \right)^{\frac{1-\alpha-\beta}{1-\beta}}}_{A_1^*} \tag{23}$$

Thus if condition (23) is met the configuration above is an equilibrium featuring no oil production and some women not working. Moreover, for any $0 \leq A_o < A_1^*$ increases in A_o have no effect on any of the equilibrium quantities or prices (we have the exact same equilibrium in this range).

A.2 Oil Production Begins

For

$$A_o > \alpha(A_c)^{\frac{1}{1-\beta}} \frac{p_c}{p_o} \left[\frac{\beta\gamma[(1-\gamma)A_s]^{\frac{1-\gamma}{\gamma}}}{\bar{u}^{\frac{1}{\gamma}}} \right]^{\frac{\beta}{1-\beta}} \left(\frac{\bar{N}_c}{L_M} \right)^{\frac{1-\alpha-\beta}{1-\beta}}$$

Oil production begins. However, given the continuity of our set-up, we expect that in the neighborhood of the threshold it should still be the case that some women don't work, whichever the effect of increased A_o is on female labor force participation. We thus search for equilibria where $A_o > A_1^*$ and (17) binds.

In such an equilibrium we have

$$w_M^* = A_o p_o \tag{24}$$
$$w_W^* = \frac{p_c \bar{u}^{\frac{1}{\gamma}}}{\gamma[(1-\gamma)A_s]^{\frac{1-\gamma}{\gamma}}}$$

Plugging these last two equations into (13) and solving for r_c yields:

$$r_c^* = (1-\alpha-\beta)(A_c p_c)^{\frac{1}{1-\alpha-\beta}} \left(\frac{\alpha}{A_o p_o}\right)^{\frac{\alpha}{1-\alpha-\beta}} \left[\frac{\beta\gamma[(1-\gamma)A_s]^{\frac{1-\gamma}{\gamma}}}{p_c \bar{u}^{\frac{1}{\gamma}}}\right]^{\frac{\beta}{1-\alpha-\beta}} \tag{25}$$

Plugging (20) and (25) into (12) and solving for L_W^c yields:

$$L_W^{c*} = (A_c p_c)^{\frac{1}{1-\alpha-\beta}} \left(\frac{\alpha}{A_o p_o}\right)^{\frac{\alpha}{1-\alpha-\beta}} \left[\frac{\beta\gamma[(1-\gamma)A_s]^{\frac{1-\gamma}{\gamma}}}{p_c \bar{u}^{\frac{1}{\gamma}}}\right]^{\frac{1-\alpha}{1-\alpha-\beta}} \tag{26}$$

Plugging (20), (24), (25) and (26) into (16), noting (9) and solving for L_W^{LF} yields:

$$L_W^{LF*} = \frac{w_W^* L_W^{c*} + (1-\gamma)A_o p_o L_M + (1-\gamma) r_c^* \bar{N}_c}{\gamma w_W^*} \tag{27}$$

Differentiating L_W^{LF*} with respect to A_o and signing the derivative yields:

$$\frac{\partial L_W^{LF*}}{\partial A_o} < 0 \text{ if}$$
$$A_o < \underbrace{\frac{\alpha(A_c p_c)^{\frac{1}{1-\beta}} \beta^{\frac{1-\alpha}{1-\beta}} [\beta + (1-\gamma)(1-\alpha-\beta)\bar{N}_c]^{\frac{1-\alpha-\beta}{1-\beta}}}{[\beta(1-\alpha-\beta)(1-\gamma)L_M]^{\frac{1-\alpha-\beta}{1-\beta}} p_o {w_W^*}^{\frac{\beta}{1-\beta}}}}_{A_2^*} \tag{28}$$

Conversely, we have that

$$\frac{\partial L_W^{LF*}}{\partial A_o} > 0 \text{ if}$$
$$A_o > \underbrace{\frac{\alpha(A_c p_c)^{\frac{1}{1-\beta}} \beta^{\frac{1-\alpha}{1-\beta}} [\beta + (1-\gamma)(1-\alpha-\beta)\bar{N}_c]^{\frac{1-\alpha-\beta}{1-\beta}}}{[\beta(1-\alpha-\beta)(1-\gamma)L_M]^{\frac{1-\alpha-\beta}{1-\beta}} p_o {w_W^*}^{\frac{\beta}{1-\beta}}}}_{A_2^*} \tag{29}$$

Two observations are in order at this stage. One is that (17) is guaranteed to bind between A_1^* and A_2^*. The imposition of restriction (5) guarantees that (17) binds at A_1^* and in the neighborhood (via continuity). Moreover between A_1^* and A_2^* increases in A_o are associated with declines in female labor force participation, such that if (17) binds at A_1^* it will also bind at A_2^* and all the intermediate points between. The second observation is that as the size of the service sector gets arbitrarily small ($\gamma \to 1$) the magnitude of A_2^* goes to infinity, which shows that in the absence of the services sector, any increase in oil sector productivity beyond the threshold at which oil production begins brings about a decline in female labor force participation (the result of Corollary 1).

A.3 A Sufficiently Large Oil Boom Brings All Women into the Labor Force

All that is left to show is that there exists and $A_3^* > A_2^*$ such that beyond A_3^* all women are drawn into work. We will prove this by contradiction. Assume no such A_3^* exists. That means that for any $A_3^* > A_2^*$ we have that (17) binds. Because $A_3^* > A_2^* > A_1^*$ oil is produced in equilibrium, which means that equations (20) and (24) to (27) hold. For $A_3^* > A_2^*$ we know that female labor force partici-

pation is increasing. We then proceed to take the limit of the number of women in the labor force, given by (27) when A_3^* goes to infinity. We are allowed to do this because by assumption (27) holds for any A_3^*. But from (27) we have that:

$$\lim_{A_3^* \to \infty} L_W^{LF*} = \infty \tag{30}$$

But by assumption, the female population of the county (L_W) is finite. We have reached a contradiction which concludes the proof.

B Appendix B: Further results and robustness checks

In this appendix, we present several additional findings and several checks of the robustness of our key findings. One first additional investigation concerns the exact nature of the service sector increase from which females are benefitting. We have shown that the share of females employed in the service sector increases. This could happen due to a variety of reasons: As manufacturing, refining and other capital-intensive sectors grow, demand for more advanced services such as banking, insurance, and business services like accounting and advertising might grow. On the other hand, income effects could lead to growth in the demand for personal services and entertainment. Of course, the types of jobs that would arise would be quite different in both scenarios. Our baseline classification includes all service sector industries and thus cannot inform us of which type of services experience particular growth. In table A1, we therefore disaggregate the service sector into three narrower categories: Personal services and Entertainment (this includes for example individuals working in

private households, in hotels, in laundering, cleaning or dyeing services and in entertainment venues), Finance, Business, and Repair Services (e.g. banking, insurance, advertising, accounting, auto repair services), and professional services (e.g. medical, legal and educational services). Overall, we had found in section 4 that the share of women employed in services increases by 5.2 percentage points after oil discoveries. As our results in table A1 show, the lion's share of this increase - 4.5 percentage points or more than 85% - comes from the personal services and entertainment category, whereas the other two categories display only very small increases. This suggests that the main reason for the blossoming of service sector jobs are income effects and the resulting increased demand for personal and recreational services. Inter-industry demand linkages that would lead to growth in professional and business services seem to be, at least in terms of female employment, less relevant.

In addition, in this section we also implement a number of robustness checks that aim to provide further support for the notion that the effects we find are indeed caused by the discovery of major oil fields.

As a first check, we show that before oil discoveries, oil-rich counties were not systematically different from counties without major oil deposits. To do so, we run a series of cross-sectional regressions for 1900:

$$y_c = \beta OilRich_c + X_c'\gamma + s_s + u_c \tag{31}$$

where $OilRich$ is 1 if a county sits on top of a large oil field an 0 otherwise.

As can be seen in table A2, across 10 different outcome variables, we do not find substantial differences between the two groups of counties. Oil- and non-

oil counties have essentially the same male and female labor force participation rate, the same distribution of the labor force over our 4 sectors of interest, and very similarly-sized populations. The only statistically significant difference arises in occupational earnings scores, where it seems that oil-rich counties on average have worse-paying occupations. However, the difference is only about one and a half percentile ranks.

Overall, the two groups of counties thus look very balanced. These findings are reassuring as they imply that the oil wealth "treatment"can be seen as nearly randomly assigned and thus the setting we study constitutes a valid natural experiment and the common trend assumption needed for the validity of our difference-in-differences design is likely to hold.

To assess this key assumption more rigorously, we augment our baseline panel specification by replacing our variable of interest with a full set of leads and lags of the date of oil discovery relevant for each oil-rich county. The specification we estimate is thus of the form:

$$y_{ct} = \alpha_c + \tau_t + \sum_{j \in \{-30,-20,-10,10,20,30\}} \beta_j DiscoveredOilField_{c,t+j} + X_c' \gamma_t + u_{ct} \tag{32}$$

where the set of dummies $DiscoveredOilField_{c,t+j}$ code for whether an oil field is to be discovered $20-30$ years from period t, $10-20$ years from period t, $10-0$ years from period t or was discovered $0-10$ years prior to period t, $10-20$ years prior to period t or more than 20 years prior to period t, with the omitted reference category being represented by discoveries that occur more than 30 years after the reference period t. All the remaining variables and controls retain their meanings from specification 6 in section 3. We show graphical results from this

specification for population and the labor force shares in our four major sectors in figures 6 to 10. Results for our remaining variables of interest are shown in tables A3 to A6.

Reassuringly, our results indicate that there is no evidence that oil-rich counties display systematically different characteristics before the discovery of oil, with virtually all the leading dummies (that indicate oil discoveries in the future) having no significant impact on any of our outcome variables of interest.

The findings related to the lagging dummies (that indicate time elapsed since discovery) largely confirm the results from our main specifications in the previous section: While there is no effect on either male or female labor force participation rates, sectoral reallocation is active. An interesting, but also intuitive further result is that the reallocation of male labor into the oil industry takes place immediately, while the increase in manufacturing needs a bit longer to materialize. In the case of the importance of services for female employment, it seems that oil counties originally display lover values and then catch up after oil discoveries take place. Yet another pattern is displayed by population, which increases sharply after oil discoveries and then continues to grow. All in all, our results from the lead and lag analysis are consistent with the results emerging from our main specifications. They support the view that the systematic differences we observe between oil-rich and baseline counties do indeed appear after oil discoveries and can be attributed to oil abundance in a causal way.

A second, related robustness check that also probes our identifying assumption involves performing our analysis on a limited sample from which all counties without oil deposits have been dropped. This is a further guard against the possibility that oil and non-oil counties might have been on different trends even

in the absence of oil discoveries. It is a very demanding empirical exercise, as it involves deriving identification only from the time variation in oil field discoveries, as well as dropping more than three quarters of our observations. We do it for our key variables of interest- employment and labor force participation rates as well as the sectoral composition of the labor force and occupational earnings scores by gender, the female population share and the share of women aged 25 and below who are married. Results are shown in tables A7 and A8. Overall, the findings of this robustness test provide a convincing validation of our previous results: having a discovered oil field is found to be associated with a reallocation of men from agriculture into oil and (to a lesser extent) manufacturing, while women increasingly work in service sector jobs. Especially men experience an increase in occupational earnings scores, which can help to explain the increase in young women's marriage propensity. Overall, both male and female labor force participation rates stay unchanged. Generally, most results from our baseline specification survive unqualified, and the point estimates of the coefficients obtained by estimating over this limited sample are very similar to those obtained in the previous section. The two exceptions are the sex ratio, where we here find a small decrease in the female population share of less than one percentage point, and the female occupational earnings scores, which in this sample display a positive evolution. However, the point estimate for women is substantially below the one of males, indicating that our conclusion of an increased gender pay gap remains the same.

Another concern with our findings is the possibility of spatial spillovers. While workforces should be less mobile in our period of analysis than nowadays, there is still a possibility that workers commute short distances and that thus coun-

ties close to oil counties also get partially treated. In addition, there might be smaller oil fields that are not in our dataset close to larger ones, such that again we would observe some spatial spillovers, a problem already noted by Michaels (2011). To address this concern of spatial spillovers, we repeat our basic specification from section 4, but drop all county x year cells pertaining to counties which border a county with a discovered oil field, but have not (yet) discovered an oil field of their own in that given year. The results of this exercise are shown in tables A9 and A10. Again, our results from before are confirmed; if anything, we find slightly larger point estimates, consistent with minor spatial spillovers. Our basic conclusions remain the same as before: Oil does not change female labor force participation, but relatively more women move from agriculture into services, and younger women are more likely to get married. Men, on the other hand, experience substantial increases in earnings that are due to sectoral reallocation from agriculture to oil mining and manufacturing.

So far, all of our robustness checks have dealt with our panel difference-in-differences specification. In table A11, we perform two robustness checks for the 1940 cross-sectional specification which we had to employ for our wage regressions. The first (columns 1 and 3) involves adding an additional dummy for "having an undiscovered oil field"to our specification. With this we are testing whether oil-rich counties in general appear to be different in the 1940 cross-section. Naturally, oil that is not yet discovered should not affect economic outcomes, so if we observe any sizeable coefficients on this dummy, we would have to be worried. Fortunately for the credibility of our cross-sectional results, this is not the case: including a dummy for an undiscovered oil field barely changes the point estimate on our variable of interest, and the estimated coefficient for

undiscovered oil fields is usually insignificant, albeit in the case of women (column 3) not very precisely estimated. In columns 2 and 4, we perform a placebo test: We drop all counties that had discovered oil fields in 1940 and compare the remaining oil-rich counties (whose oil wealth had not yet been discovered) to counties without oil. Again, if our identifying assumptions are valid, we would expect to see no effect of the undiscovered oil field dummy, and again this is the case, which provides further reassurance for our cross-sectional estimates.

	(1)	(2)	(3)
VARIABLES	Share of the female LF employed in		
	Personal serv. and entertainment	Finance, Business, and repair serv.	Professional serv.
Discovered Oil Field	0.045	0.004	0.003
	(0.027)	(0.003)	(0.006)
Mean Dep Var	0.331	0.024	0.125

Number of observations 3,475, number of clusters 774. Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** p<0.01, ** p<0.05, * p<0.1

Table A1: Dissecting the female service sector increase

VARIABLES	(1) Labor Force Participation Rate Female	(2) Male	(3) Occ. earnings score Female	(4) Male	(5) Urban Pop Share
Oil Rich	0.001	0.004	-1.674*	-1.569	-0.071
	(0.016)	(0.005)	(0.924)	(2.444)	(0.067)
Observations	587	588	549	588	588
Mean Dep Var	0.213	0.930	15.121	27.509	0.233

VARIABLES	(6) Share of the Labor Force employed in Oil mining	(7) Manufacturing	(8) Services	(9) Agriculture	(10) ln(pop)
Oil Rich	-0.000	0.004	-0.008	0.026	-0.023
	(0.000)	(0.008)	(0.021)	(0.055)	(0.102)
Observations	588	588	588	588	588
Mean Dep Var	0.000	0.051	0.128	0.63	

Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** p<0.01, ** p<0.05, * p<0.1

Table A2: Balancing checks in 1900

VARIABLES	(1) Female pop share	(2) LF Part Rate	(3) Employment Rate	(4) Average occ. earnings score
-30 to -21	0.000	0.009	0.004	0.360
	(0.003)	(0.006)	(0.008)	(0.760)
-20 to -11	-0.000	0.012*	0.018	-0.759
	(0.004)	(0.007)	(0.011)	(0.809)
-10 to -1	0.004	0.017*	0.031***	-1.158
	(0.004)	(0.009)	(0.011)	(0.950)
0 to 10	-0.005	0.020*	0.018	3.055***
	(0.005)	(0.010)	(0.011)	(1.058)
11 to 20	0.005	0.026**	0.014	4.198***
	(0.006)	(0.013)	(0.012)	(1.507)
20+	0.008	0.005	0.019	3.717**
	(0.008)	(0.034)	(0.013)	(1.512)
Sample	All	Men	Male LF	Male LF
Observations	3,594	3,594	2,231	3,577

Number of clusters 774. Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)
*** p<0.01, ** p<0.05, * p<0.1

Table A3: Leads and lags analysis, part 1

	(1)	(2)	(3)	(4)
VARIABLES	Share of Male LF employed in			
	Oil Mining	Manufacturing	Services	Agriculture
-30 to -21	0.003	0.007	0.006	0.006
	(0.003)	(0.006)	(0.005)	(0.013)
-20 to -11	-0.002	-0.004	-0.001	0.026*
	(0.004)	(0.010)	(0.004)	(0.015)
-10 to -1	-0.006	0.004	-0.002	0.039**
	(0.005)	(0.012)	(0.005)	(0.019)
0 to 10	0.061***	0.008	0.001	-0.049**
	(0.010)	(0.011)	(0.005)	(0.021)
11 to 20	0.064***	0.021*	0.008	-0.077***
	(0.090)	(0.012)	(0.006)	(0.023)
20+	0.043***	0.033**	0.012	-0.069**
	(0.090)	(0.014)	(0.090)	(0.028)

Number of observations 3,577, number of clusters 774. Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** p<0.01, ** p<0.05, * p<0.1

Table A4: Leads and lags analysis, part 2

VARIABLES	(1) LF Part Rate	(2) Share Married	(3) Employment Rate	(4) Average occ.
-30 to -21	-0.016	0.006	0.006	0.361
	(0.018)	(0.011)	(0.010)	(0.664)
-20 to -11	-0.008	0.011	-0.010	0.448
	(0.025)	(0.013)	(0.015)	(0.730)
-10 to -1	-0.005	0.008	0.015	0.281
	(0.025)	(0.015)	(0.015)	(0.868)
0 to 10	-0.015	0.026*	0.000	1.073
	(0.026)	(0.015)	(0.016)	(0.843)
11 to 20	-0.005	0.019	-0.008	0.745
	(0.028)	(0.017)	(0.016)	(1.058)
20+	0.001	0.024	0.010	-0.295
	(0.033)	(0.024)	(0.018)	(1.147)
Sample	All Women	Women ≤ 25	Female LF	Female LF
Observations	3,590	3,573	2,171	3,478

Number of clusters 774. Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** p<0.01, ** p<0.05, * p<0.1

Table A5: Leads and lags analysis, part 3

	(1)	(2)	(3)	(4)
VARIABLES	Share of Female LF employed in			
	Oil Mining	Manufacturing	Services	Agriculture
-30 to -21	-0.000	0.001	-0.029	0.023
	(0.000)	(0.006)	(0.055)	(0.049)
-20 to -11	-0.000	-0.006	-0.047	0.048
	(0.000)	(0.007)	(0.079)	(0.072)
-10 to -1	-0.000	0.001	-0.043	0.035
	(0.000)	(0.008)	(0.072)	(0.065)
0 to 10	0.001***	-0.008	-0.004	-0.015
	(0.000)	(0.008)	(0.078)	(0.071)
11 to 20	0.001	-0.013	0.032	-0.045
	(0.001)	(0.010)	(0.085)	(0.079)
20+	0.004***	-0.019	-0.006	0.001
	(0.001)	(0.013)	(0.101)	(0.093)

Number of observations 3,475, number of clusters 774. Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** p<0.01, ** p<0.05, * p<0.1

Table A6: Leads and lags analysis, part 4

VARIABLES	(1) Female pop share	(2) LF Part Rate	(3) Employment Rate	(4) Average occ. earnings score
Disc Oil Field	-0.007**	-0.001	-0.016**	3.523***
	(0.003)	(0.007)	(0.007)	(0.941)
Sample	All	Men	Male LF	Male LF
Observations	771	771	487	766
Clusters	171	171	171	171

VARIABLES	(5)	(6)	(7)	(8)
	Share of Male LF employed in			
	Oil Mining	Manufacturing	Services	Agriculture
Disc Oil Field	0.063***	0.005	0.003	-0.074***
	(0.010)	(0.005)	(0.004)	(0.017)

Number of observations 766, number of clusters 171. Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** p<0.01, ** p<0.05, * p<0.1

Table A7: Results when dropping all counties without an oil deposit, part 1

VARIABLES	(1) LF Part Rate	(2) Share Married	(3) Employment Rate	(4) Average occ. earnings score
Disc Oil Field	-0.005	0.014	-0.009	1.148*
	(0.011)	(0.011)	(0.009)	(0.638)
Sample	All Women	Women ≤ 25	Female LF	Female LF
Observations	770	765	464	733
Clusters	171	171	171	171

	(5)	(6)	(7)	(8)
VARIABLES		Share of Female LF employed in		
	Oil Mining	Manufacturing	Services	Agriculture
Disc Oil Field	0.000	-0.005	0.036	-0.050
	(0.001)	(0.005)	(0.028)	(0.031)

Number of observations 733, number of clusters 171. Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** p<0.01, ** p<0.05, * p<0.1

Table A8: Results when dropping all counties without an oil deposit, part 2

	(1)	(2)	(3)	(4)
VARIABLES	Female pop share	LF Part Rate	Employment Rate	Average occ. earnings score
Disc Oil Field	-0.006	0.013*	-0.001	4.277***
	(0.003)	(0.007)	(0.005)	(1.042)
Sample	All	Men	Male LF	Male LF
Observations	3,027	3,027	1,775	3,015
Clusters	741	741	722	741

	(5)	(6)	(7)	(8)
VARIABLES		Share of Male LF employed in		
	Oil Mining	Manufacturing	Services	Agriculture
Disc Oil Field	0.071***	0.010**	0.002	-0.086***
	(0.010)	(0.004)	(0.004)	(0.018)

Number of observations 3,015, number of clusters 741. Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** p<0.01, ** p<0.05, * p<0.1

Table A9: Results when excluding neighboring counties without discovered oil wealth, part 1

VARIABLES	(1) LF Part Rate	(2) Share Married	(3) Employment Rate	(4) Average occ. earnings score
Disc Oil Field	-0.006	0.022**	0.004	0.424
	(0.012)	(0.010)	(0.011)	(0.757)
Sample	All Women	Women ≤ 25	Female LF	Female LF
Observations	3,023	3,007	1,717	2,918
Clusters	741	740	710	738

VARIABLES	(5) Oil Mining	(6) Share of Female LF employed in Manufacturing	(7) Services	(8) Agriculture
Disc Oil Field	0.001***	-0.010*	0.052*	-0.062**
	(0.000)	(0.006)	(0.027)	(0.028)

Number of observations 2,915, number of clusters 738. Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** p<0.01, ** p<0.05, * p<0.1

Table A10: Results when excluding neighboring counties without discovered oil wealth, part 2

	(1)	(2)	(3)	(4)
VARIABLES	Average ln(annual wage income)			
Discovered Oil Field	0.258***		0.038	
	(0.083)		(0.072)	
Not Yet Discovered	-0.042	-0.034	-0.072	-0.054
	(0.102)	(0.104)	(0.090)	(0.0903)
Counties	All	Omit disc oil fields	All	Omit disc oil fields
Sample	Male LF	Male LF	Female LF	Female LF
Observations	758	643	758	643

Standard errors, clustered at the county level, in parentheses. Standard errors and p-values are based on bias adjusted standard errors and effective degrees of freedom corrections as in Young (2016)

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A11: Robustness checks 1940 cross section

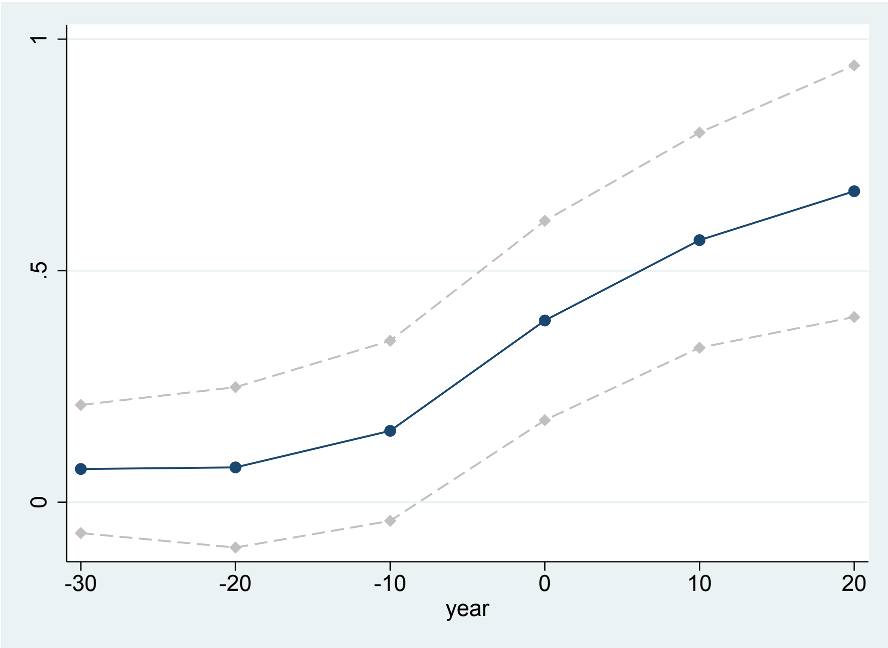

Figure 6: Leads and lags for log population

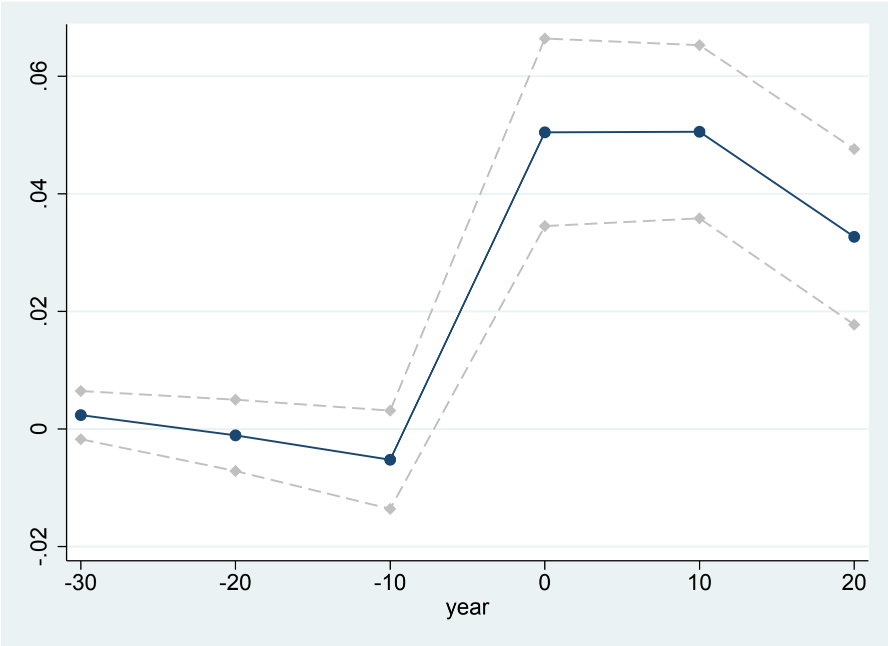

Figure 7: Leads and lags for the LF share employed in oil mining

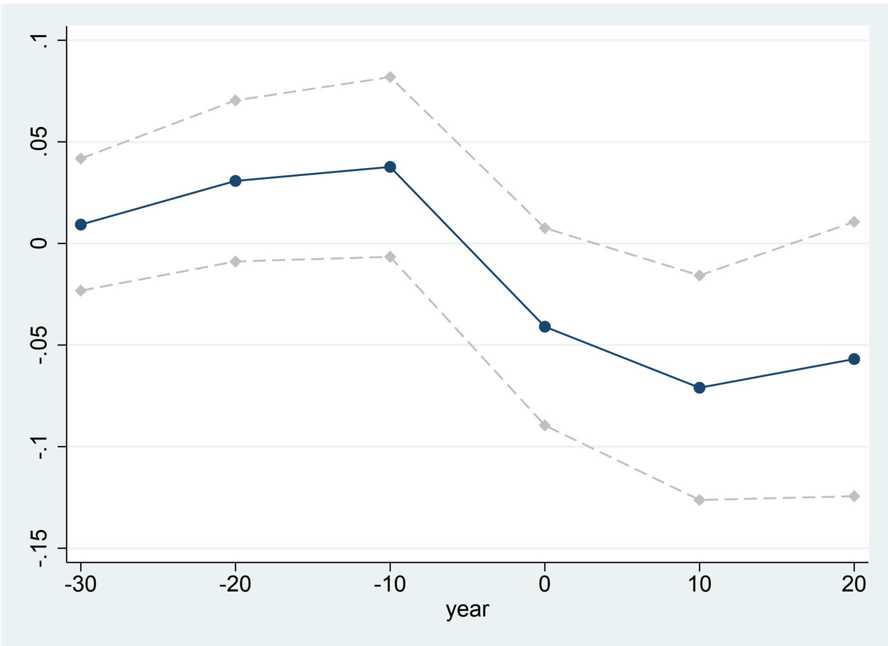

Figure 8: Leads and lags for the LF share employed in agriculture

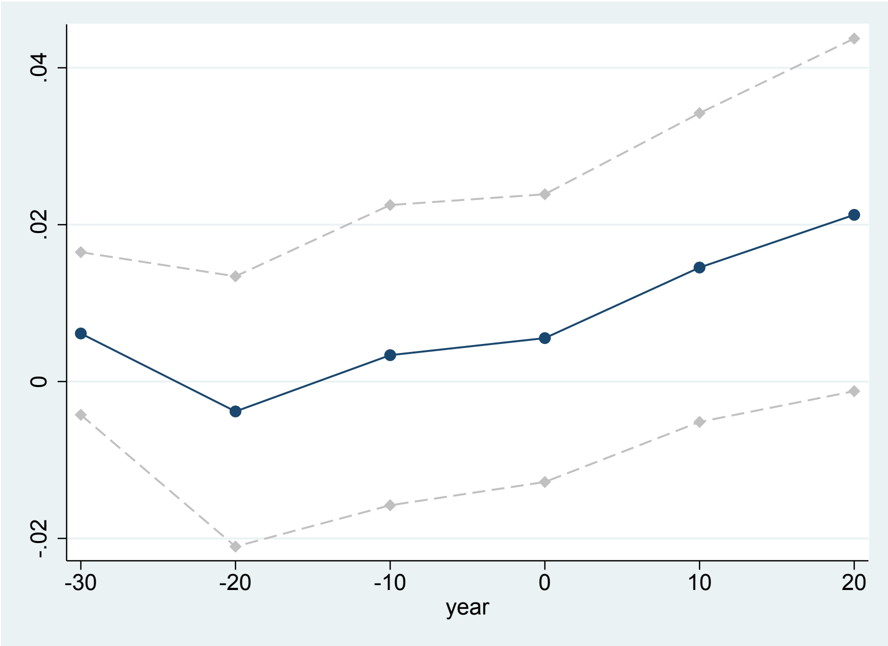

Figure 9: Leads and lags for the LF share employed in manufacturing

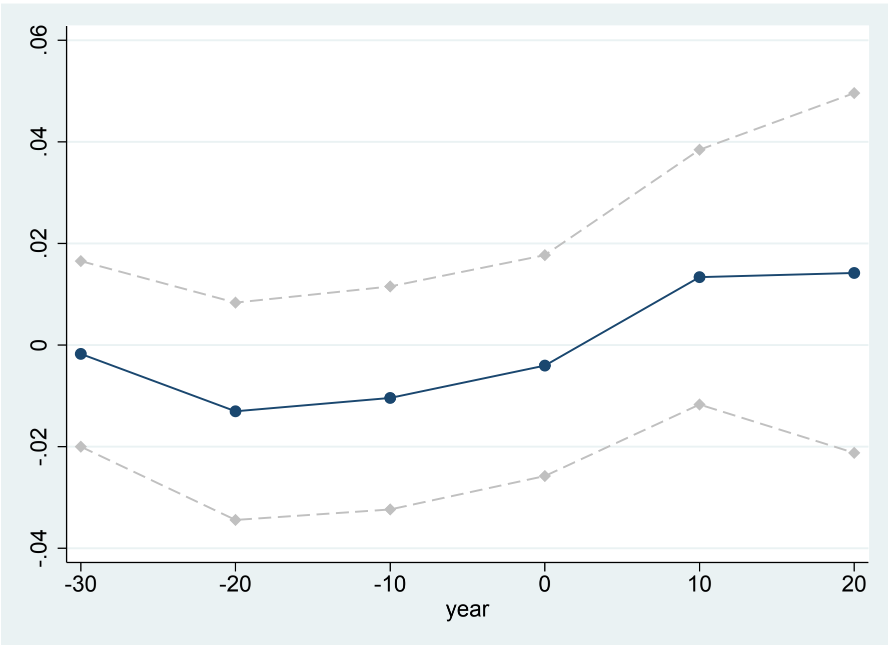

Figure 10: Leads and lags for the LF share employed in services