Gender Difference in Labor Demand Behavior

Oluwasheyi Oladipo* Kasia Platt† Hyoung Suk Shim‡

December 31, 2019

Work in progress, Comments welcome

Abstract

This paper empirically examines firm owners' gender difference in labor demand. We estimate the average treatment effect (ATE) of female ownership on employment of the firm using the 2007 Survey of Business Owners (SBO), provided by the Census Bureau. We show first that the estimation of the ATE for female ownership can have a downward selection bias that causes negative ATE estimate, and this downward selection bias comes from difference in financial constraint by firm owners' gender in which male owners are less financially constrained than female owners. We then perform the IV estimation and the two-stage least squares (TSLS) using indicator variables for start-up capital formation, bank loan and family/friend loan, and Inheritance as two different sets of IVs. The estimation results present that the female owner effect on labor demand as local average treatment effect (LATE) is identified and consistently estimated by using the IVs. From the main model estimation, we find that a positive and statistically significant female owner effect that female owners hire more employees than male owner by about 25.8%.

1 Identification

We statistically identify the female owner effect on labor demand as local average treatment effect (LATE), and estimate it using the IV estimation and the two-stage least squares (TSLS). Simply, LATE is an instrumental variable (IV) estimation of ATE with binary IVs.

*Department of Politics, Economics & Law, State University of New York at Old Westbury, Long Island, New York, email: oladipoS@oldwestbury.edu.

†School of Business, State University of New York at Old Westbury, Long Island, New York, email: plattk@oldwestbury.edu.

‡School of Business, College of Staten Island, The City University of New York, New York, email: Hyoungsuk.Shim@csi.cuny.edu.

Consider the cost-minimization of a firm, as in Hamermesh (1996) to derive labor demand, denoted L. An observable form of the labor demand function is $L^* = L^d(w, r, Y)$, where w is wage, r is interest, and Y is output level. Consider a binary indicator variable $D_i = \{0, 1\}$ for the gender of firm i's owner. We use D_i as a female owner indicator so $D_i = 1$ if firm i's owner is female. For any firm, there are two potential labor demand variables:

$$\begin{aligned} L_i &= \begin{cases} L_{1i} & \text{if } D_i = 1 \\ L_{0i} & \text{if } D_i = 0 \end{cases} \\ &= L_{0i} + (L_{1i} - L_{0i}) \cdot D_i. \end{aligned} \tag{1.1}$$

The observable labor demand for firm i consists of two potential labor demand. That is L_{1i} if firm i is managed by a female owner, and L_{0i} if the manager is male. Our causal effect of interest is $L_{1i} - L_{0i}$, the difference in labor demand by gender, but it is not directly observable. What we can observe with L_i and D_i is

$$\begin{aligned} \underbrace{\mathrm{E}[L_i|D_i = 1] - \mathrm{E}[L_i|D_i = 0]}_{\text{Observed difference in average}} &= \underbrace{\mathrm{E}[L_{1i}|D_i = 1] - \mathrm{E}[L_{0i}|D_i = 1]}_{\text{average treatment effect on the treated}} \\ &\quad + \underbrace{\mathrm{E}[L_{0i}|D_i = 1] - \mathrm{E}[L_{0i}|D_i = 0]}_{\text{selection bias}}. \end{aligned}$$

We argue that the selection bias is negative because male owners are less financially constrained, and thus, for given wage w and output Y,

$$L_{1i}(w, r_1, Y) \leq L_{0i}(w, r_0, Y) \tag{1.2}$$

where r_1 and r_0 are interest cost for female and male owner firms respectively. Less financially constrained male owner would have less interest cost so that $r_1 \geq r_0$. The implication of the term $\mathrm{E}[L_{0i}|D_i = 1]$ is an average labor demand for male owner firms under female owner firms' financial condition. In the same way, $\mathrm{E}[L_{0i}|D_i = 0]$ implies an average labor demand for male owner firms under male owner firms' financial condition. For financial institutions that firm i would like to borrow capital, L_{1i} and L_{0i} are unobservable so that they apply r_1, and r_0 by looking at the treatment status D_i. Therefore,

$$\begin{aligned} &\mathrm{E}[L_{0i}|D_i = 1] - \mathrm{E}[L_{0i}|D_i = 0] \\ &\qquad = \mathrm{E}[L_{0i}|D_i = 1, r = r_1] - \mathrm{E}[L_{0i}|D_i = 0, r = r_0] \\ &\qquad \leq 0 \end{aligned}$$

The interest cost assignment r_1, r_0 are useful for the LATE parameter to be identified and consistently estimated. Monotonicity is one of the four LATE assumptions to be a consistent estimator. It asserts that the treatment assignment is accepted in the same way by all individuals. That is $D_{1i} \geq D_{0i}$ or $D_{1i} \leq D_{0i}$ for all i. In our case, $D_{1i} \leq D_{0i}$ meaning that firm owners prefer to be treated as male because of financial constraints and interest cost $r_1 \geq r_0$.

We use indicator variables for start-up capital formation, bank loan and family/friend loan as IVs. Inheritance is another indicator variable that we use as an IV. It is one if a business owner has bequeathed the business as inheritance, and thus it differentiates firms with and without financial constraints. Since the firm owners receiving the businesses as inheritance have no interest cost, the inheritance IV is strongly correlated with interest cost and uncorrelated with wage level and product demand shock. We argue that the IVs are valid for following reasons: i) the start-up capital formation is correlated with interest cost r so D_i; ii) and it is uncorrelated with wage level and product demand shock.

With the IVs, the female owner effect on labor demand can be identified and consistently estimated as local average treatment effect (LATE). To show this, consider a simple linear regression model from (1.1)

$$\begin{aligned} L_i &= L_{0i} + (L_{1i} - L_{0i}) \cdot D_i \\ &= \mathrm{E}[L_{0i}] + (L_{1i} - L_{0i}) \cdot D_i + (L_{0i} - \mathrm{E}[L_{0i}]) \\ &= \alpha + \rho_i \cdot D_i + \eta_i, \end{aligned}$$

where ρ_i is a random coefficient representation of the ATE, and η_i is an error term.

LATE estimate with multiple instruments is a weighted average of Wald estimators for each instrument. In our case, the female owner effect on labor demand is estimated as the weighted average of two Wald estimates, one with bank loan indicator and the other with family/friend loan indicator. Let Z_{0i} be the inheritance indicator, and let Z_{1i} and Z_{2i} be the bank loan indicator and family/friend loan indicator variables respectively. For the inheritance IV, the LATE parameter, denote ρ as an average of the random coefficient ρ_i can be identified

$$\begin{aligned} \mathrm{E}[\rho_i | D_{1i} < D_{0i}] &= \frac{\mathrm{Cov}(L_i, Z_{0i})}{\mathrm{Cov}(D_i, Z_{0i})} \\ &= \rho \end{aligned} \tag{1.3}$$

For the ATE estimation of ρ_i, we have two IV estimands,

$$\rho_1 = \frac{\text{Cov}(L_i, Z_{1i})}{\text{Cov}(D_i, Z_{1i})}, \quad \rho_2 = \frac{\text{Cov}(L_i, Z_{2i})}{\text{Cov}(D_i, Z_{2i})}.$$

With the first-stage fitted value $\hat{D}_i = \pi_1 \cdot Z_{1i} + \pi_2 \cdot Z_{2i}$, the two-stage least squares (TSLS) estimand for ρ is then

$$\begin{aligned} \rho &= \frac{\text{Cov}(L_i, \hat{D}_i)}{\text{Cov}(D_i, \hat{D}_i)} \\ &= \psi \cdot \rho_1 + (1 - \psi)\rho_2, \end{aligned} \tag{1.4}$$

where

$$\psi = \frac{\pi_1 \cdot \text{Cov}(D_i, Z_{1i})}{\pi_1 \cdot \text{Cov}(D_i, Z_{1i}) + \pi_2 \cdot \text{Cov}(D_i, Z_{2i})},$$

is a fraction. Intuitively, the inheritance status would be a stronger IV than the loan IVs but the number of firms with inheritance in the 2007 SBO is very small. We thus estimate the female owner effect ρ using (1.3) and (1.4) with inheritance IV and loan IVs separately, and carefully examine test statistics for their endogeneity and first-stage F-test.

2 Data

We make use of the 2007 Survey of Business Owners (SBO) to create a dataset for the labor demand model estimation with owners' gender. The SBO is a 5-year period survey for operating firms and companies in the United States, conducted by The Census Bureau. Firms in the survey are randomly selected from the list of firms that filed their tax report with the Internal Revenue Service (IRS). The Census Bureau obtains the sample firms' employment numbers, payroll, and receipts from their IRS tax reports. Other information related to the firm owners' demographics and their business operations are collected via mail. There are 663,385 single owner firms from a total of 2,165,680 firm records in the 2007 SBO sample. In our dataset, about 33% of the firms are female-owned.

Table 1 reports descriptive statistics for the SBO data by firm owner's gender. Employment and payroll are records from the IRS tax reports. The statistics in table 1 are all weighted by the SBO tabulation weight. Start-up capital is originally given as a categorical variables. We calculate and report its descriptive statistics by assigning the middle value of each category. Inheritance, bank loan, and family loan are binary indicator variables to be used as IVs. The SBO has inheritance status and start-up capital formation method in its

Table 1: Descriptive Statistics by Gender and Homebase

	Female Owner					Male Owner				
	# of Firms	Mean	Std	5^{th}	95^{th}	# of Firms	Mean	Std	5^{th}	95^{th}
Employment	220,625	0.848	15.73	0	4	442,760	2.144	31.26	0	8
Start-up Capital	135,847	25,631	91,728	2,500	77,500	301,479	46,130	133,121	2,500	175,000
Payroll	220,625	21.145	430.57	0	70	442,760	70.249	864.63	0	270
Inheritance	211,872	0.011	0.10	0	0	433,136	0.010	0.10	0	0
Bank Loan	220,625	0.150	0.36	0	1	442,760	0.188	0.39	0	1
Family Loan	220,625	0.014	0.12	0	0	442,760	0.021	0.14	0	0
With Spouse	218,177	0.0393	0.194	0	0	438,123	0.0579	0.234	0	1
Family business	219,511	0.0180	0.133	0	0	440,766	0.0258	0.159	0	0
Education	215,284	4.557	1.92	2	7	431,534	4.475	2.03	1	7
Age	215,915	3.827	1.27	2	6	434,257	3.980	1.29	2	6
Nonwhite	220,625	0.136	0.34	0	1	442,760	0.102	0.30	0	1
Years of Operation	201,699	4.009	2.64	0	8	416,507	4.669	2.63	0	8

[#] The reported statistics are weighted by the SBO tabulation weight. Education is an ordinal categorical variable 1 = less than high school, 2 = high school, 3 = technical school, 4 = some college, 5 = associate degree, 6 = bachelor degree, 7 = masters or above. Age is another ordinal categorical variable: 1 = under 25, 2 = 25 to 24, 3 = 35 to 44, 4 = 45 to 54, 5 = 55 to 64, 6 = 65 or over. Years of operation is also ordinal categorical variable: 1 = from 2007, 2 = from 2006, 3 = from 2005, 4 = from 2004, 5 = from 2003, 6 = from 2000 and 2002, 7 = from 1990 and 1999, 8 = from between 1980 and 1989, 9 = from before 1980.

questionaries. We use them to create the three binary IVs. Difference in production related variables between female and male owners are clearly shown in table 1, whereas financial constraints, demographic variables are not. For male owner firms, the average employment and start-up capital size are about twice as big as female owner firms, though the standard deviations are way too big to confirm that the differences are statistically significant. Payroll expense of male owner firms, in average, are about three times bigger than female owner firms. These are weak and insignificant evidences that female owner firms are smiller than male owner firms in terms of production input, capital and labor.

Our identification strategy is to use inheritance, and loan from bank or friend/family as IVs. The inheritance IV seems to have too few treatment observations, 1.1% female owner firms and 1.0% male owner firms, and this may cause inconsistent estimation due to weak instrument. This is one of the reasons that we consider the other set of instrumental variables, loan from bank or friend/family. About 19.54% of female owners have issued loans from either bank or friend/family. The fraction for male owners is 25.27%. There might be a trade-off between which the inheritance IV has much stronger correlation with unobserved interest cost but too few observations, and the loan IVs have not much strong correlation but relatively enough observations. We therefore estimate the female owner effect with inheritance IV and loan IVs separately, and carefully examine test statistics for their endogeneity and first-stage F-test.

Table 2: Main Model Estimates: Log of Employment

	OLS			IV Estimate with Inheritance			TSLS with Loans as IVs		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Female Owner	-0.2063***	0.1522***	0.1250***	5.3204***	0.8495***	0.2295***	20.9104***	10.9855***	9.3862***
	[0.019]	[0.007]	[0.005]	[0.412]	[0.092]	[0.053]	[4.772]	[1.690]	[2.741]
Payroll		0.6938***	0.6887***		0.7128***	0.6906***		0.9836***	0.8698***
(Log of)		[0.004]	[0.004]		[0.005]	[0.004]		[0.052]	[0.049]
Control Variables	No	No	Yes	No	No	Yes	No	No	Yes
State Fixed	No	No	Yes	No	No	Yes	No	No	Yes
Industry Fixed	No	No	Yes	No	No	Yes	No	No	Yes
# of Obs	267,826	260,670	242,910	264,584	257,555	242,021	267,826	260,670	242,021
Adj R^2	0.0035	0.7138	0.7492	NA	0.6746	0.7484	NA	NA	NA
F-Test (df_n, df_d)				386.89(1,42)	52.10(1,42)	3.39(1,42)	337.75(1,42)	338.62(1,42)	126.67(1,42)
(P-value)				(0.0000)	(0.0000)	(0.0726)	(0.0000)	(0.0000)	(0.0000)
F-stat {First-stage}	123.65	17,297.76	NA	{162.08}	{204.08}	{326.65}	{11.78}	{22.41}	{6.83}

Notes: Heteroskedasticity robust standard errors clustered by state are reported in square brackets. The symbols, *, **, and *** indicate respectively that the estimated coefficient is statistically significant under 10%, 5%, and 1% significance levels.

3 Empirical Results

The main model estimates are reported in table 2. We find that the model estimation results are consistent with our prediction on the selection bias. The first column of each three panel are the ATE estimates without any control variables. The two IV estimates in the middle and right panels are positive and strongly significant, whereas the OLS estimate in the left panel is negative with strong significance. This downward bias of the OLS estimation is observed in the rest of the two model estimates consistently. The OLS estimates in the second and third columns turn out to be positive with strong statistical significance. However, it is smaller than the two corresponding IV estimates. The full model estimates in the third column have the same pattern that the OLS estimate is smaller than the other two IV estimates.

The inheritance IV seems to perform better than the two loan IVs in terms of controlling endogenous female effect due to the selection bias, and not being a weak instrument. The two different IV estimations, reported in the middle and right panels in table 2, are all positive with strong statistical significance, but the TSLS estimates with the loan IVs are much bigger than the IV estimates with inheritance. This difference may come from the bias caused by weak instruments. The first-stage F-test statistics for the IV estimations are reported in the last row in table 2. The well known threshold for the F-test statistic not causing bias from weak instrument is above 10, as suggested by Stock et al. (2002). The three IV estimates in the middle panel have the F-stat way above 100, while the three TSLS estimates in the right panel have F-stat around 10. Since the three TSLS estimates are all

greater than the corresponding IV estimates, the loan IVs cause upward bias due to weak correlation with the female owner treatment indicator.

The log of payroll variable seems to be an effective control for the endogenous female owner effect. The second and third columns in the OLS panel of table 2 report positive and reasonable size of the female owner effect estimates. The six IV estimates present similar patterns around the payroll variable. This is also consistent with our prediction about the selection bias in the female owner effect on labor demand. The source of bias is endogenous interest cost assignment between female and male owners. The interest cost for a firm is not directly observable, but it affects the optimal factor (labor) demand for the firm. Therefore, the selection bias can be mitigated substantially by controlling for the firm's expenditure on labor. Still, the female owner effect estimate from OLS with the full model specification, reported in the third column of the left panel in table 2, is about half of the full model IV estimate reported in the middle panel with which the endogenous IV test confirms its validity under 5% significance level. From this, the most reliable female owner effect is 0.2295 from the IV estimation with the full model specification, and this implies that female owner hires more employees than male owner by about 25.8%.

We then estimate a model for probability of being employer. This model specification is similar to Fairlie and Miranda (2017), where the gender effect on probability of hiring the first employee is estimated. They estimate that female-owned firms are about 10% point less likely to hire their first employee. Another benefit of the probability model estimation is to consider self-employed or nonemployer firms where the number of employments is zero. These firms are omitted from the model estimations in table 2. and this may cause the opposity way of selection bias. We can thus check whether the female owner effect estimation is affected by the omitted observations, and compare to the result of Fairlie and Miranda (2017).

Estimation results for the probability model are reported in table 3. As shown in table 2, the negative female owner effect seems to be a result of selection bias due to the endogenous interest cost assignment. We estimate the model using four different estimations, and find that the female owner effect is negative in non-IV estimations. But, it becomes positive with strong significance in IV estimations. In table 3, the first two panels are the non-IV estimates. The OLS estimations with linear probability model (LPM) yield a female owner effect of about around -0.2 and the maximum likelihood estimations (MLE) with probit model specification result the female owner effect of about around -0.6. The two panels on the right in table 3 are IV estimation results. All of four female owner effects are positive with strong statistical significance, but the sizes differ by the presence of control variables. The IV estimations with LPM specification yield 8.31 without control variables, and 0.751 with control variables as the female owner effect estimates. In the same way, the two-step

Table 3: Main Model Estimates: Probability of being Employer

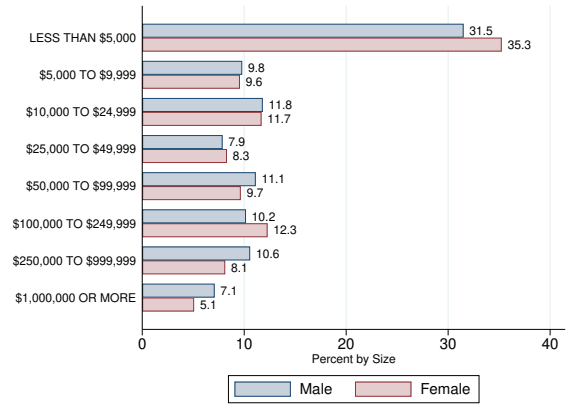
	OLS		Probit		IV Estimation with Inheritance		IV Probit with Inheritance	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Female Owner	-0.2534***	-0.1835***	-0.6845***	-0.5477***	8.3079***	0.7510***	23.1498***	2.2100***
	[0.004]	[0.004]	[0.013]	[0.012]	[2.774]	[0.096]	[3.487]	[0.149]
Control Variables	No	Yes	No	Yes	No	Yes	No	Yes
State Fixed	No	Yes	No	Yes	No	Yes	No	Yes
Industry Fixed	No	Yes	No	Yes	No	Yes	No	Yes
# of Obs	642,194	571,651	642,194	571,651	624,334	571,651	645,008	571,651
Adj R^2	0.0585	0.1900	NA	NA	NA	NA	NA	NA
χ^2 Test (df)					732.82	281.06	3371.02(1)	603.12(1)
(P-value)					(0.0000)	(0.0000)	(0.0000)	(0.0000)
F-stat {First-stage}	3,223.91	NA			{10.05}	{191.39}		

Notes: Heteroskedasticity robust standard errors clustered by state are reported in square brackets. The symbols, *, **, and *** indicate respectively that the estimated coefficient is statistically significant under 10%, 5%, and 1% significance levels.

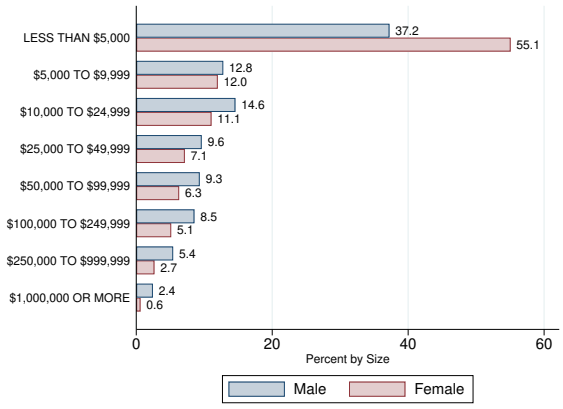
MLE with probit specification estimate 23.15 and 2.21 with and without control variables respectively.

Overall, we can see that there is a downward bias in estimating female owner effect on probability of being employer. The negative effect is thus a result of the bias, and the main source seems to be the endogenous interest cost assignment between female and male owners. The evidence for this is that the inheritance IV estimates a positive female owner effect as in table 2. Obtaining a precise estimate of the effect is, however, invalid with our dataset, since there is a control for production or labor cost.

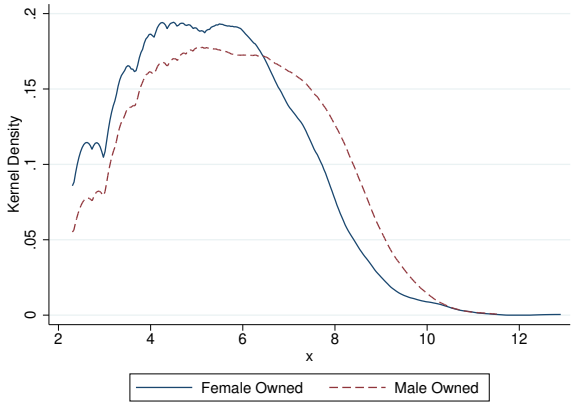
We estimate distributions of factor cost expenditure by gender and inheritance status. This is to examine whether there is a behavioral difference in factor demand and expenditure between female and male owners, other than the difference in labor demand. The distribution of start-up capital size by gender and inheritance status are presented on the top of figure 1. Firms with inheritance have an almost identical distribution of startup capital between female and male owners, whereas firms without inheritance have smaller start-up capital for female than male owners. This is a descriptive evidence that female owners are more likely to be financially constrained, and inheritance status is a valid IV that can rule out the difference in financial constraint. The bottom two panels in figure 1 are nonparametric kernel distributions of payroll. Overall, they present that male owners have a little mire labor expenditure but the difference with female owners does not look substantially big. Inheritance status seems to change the overall shape of the payroll distributions but the



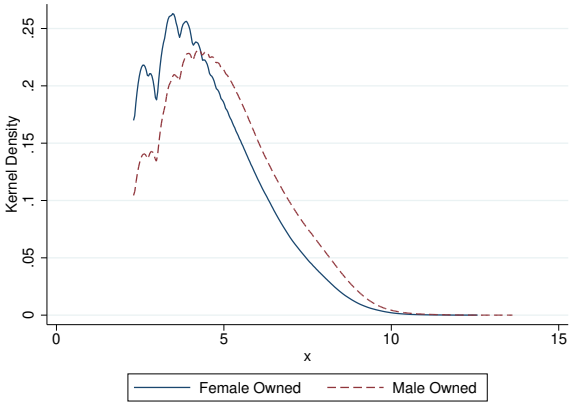
(a) Start-up Capital with Inheritance



(b) Start-up Capital without Inheritance



(c) (Log of) Payroll with Inheritance



(d) (Log of) Payroll without Inheritance

Figure 1: Factor Cost Distribution by Credit Constraint

Table 4: Model Estimates: Log of Employment by Factor Demand

	Homebase Firms		Firms with Spouse		Firms with Family	
	Homebased	Non-Homebased	With Spouse	Without Spouse	Family Businss	Non-family Businss
Female Owner	-0.0813	0.1694***	-0.0681	0.2387***	0.1595	0.2359***
	[0.127]	[0.053]	[0.186]	[0.053]	[0.134]	[0.054]
Payroll (Log of)	0.4403***	0.7147***	0.6696***	0.6915***	0.7082***	0.6900***
	[0.009]	[0.003]	[0.007]	[0.004]	[0.004]	[0.004]
State Fixed	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs	48,890	194,900	11,400	231,372	8,073	235,721
Adj R^2	0.4023	0.7610	0.7101	0.7498	0.7560	0.7479
F-Test (df_n, df_d)	2.25(1,42)	0.38(1,42)	0.86(1,42)	4.01(1,42)	0.06(1,42)	3.78(1,42)
(P-value)	(0.1410)	(0.5406)	(0.3604)	(0.0516)	(0.8053)	(0.0586)
F-stat {First-stage}	{98.51}	{331.33}	{35.80}	{293.06}	{137.18}	{299.36}

Notes: Heteroskedasticity robust standard errors clustered by state are reported in square brackets. The symbols, *, **, and *** indicate respectively that the estimated coefficient is statistically significant under 10%, 5%, and 1% significance levels.

difference does not change.

3.1 Demand for Flexible Work

The positive female owner effect cannot be explained by the endogenous interest cost assignment alone. Rather, the effect would be insignificant, since firms seek to hire optimal number of employees for profit maximization, and the optimum cannot be different by owners' gender. We thus empirically examine the role of demand for flexible work as a possible channel through which female owners are likely to demand more employees. In labor demand literature, female workers' preference for flexible work has been discussed in a number of papers such as Wiswall and Zafar (2018), but these have no emphasis on labor demand.

Table 4 reports IV estimates of the female owner effect by six subsets for different household labor demand condition. The left panel reports the estimated female owner effect by homebased business, and shows that non-homebased business has a positive female owner effect with strong statistical significance, whereas home based business is insignificant. The rest of model estimates in the middle and right panels have the same pattern as the homebase subset estimates. The female owner effect is positive with strong statistical significance, if the owner runs the business with husband or family. And, if not, the female owner effect becomes insignificant. For these six model estimates, the inheritance IV works well to control for the endogeneity without weak instrument bias. The endogeneity F-stats yield p-values greater than 0.05, and the first-stage F-stats are well above 10.

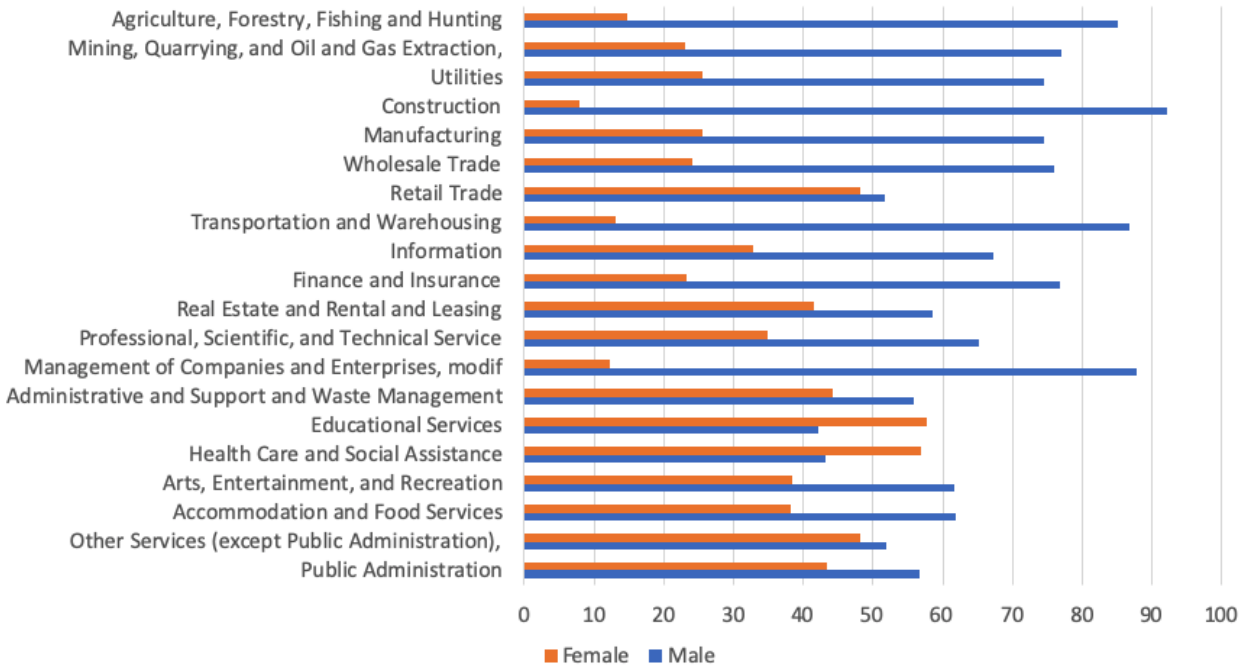


Figure 2: Fraction of Business Owners' Gender by Industry

We find an indirect evidence that the positive female owner effect comes from family issues such as marriage, household production, and child cares. In table 4, the female owner effects are insignificant for home-based, and businesses with spouse or family. These are the condition under which female owners can spend less time and cost for the household production. On the other hand, the female owner effects are positive with strong statistical significance for non-home based, or without spouse and family. A number of papers on female labor supply have discussed about the effect of family factors. Especially, our finding is consistent with Edwards and Field-Hendrey (2002) that female labor force willing to lower "the fixed costs of working (e.g., time costs associated with commuting, out-of-pocket commuting expenditures, and clothing costs)", which imply that they have spent additional cost to allocate more time for household production and other family matters.

3.2 Robustness Check

For checking robustness, we separately estimate the female onwer effect by industry, size of start-up capital, and firm age. de Mel et al. (2009) report that industries with more female owners have less investment rates and returns to investment. In other words, female owner oriented industries are either less focusing on its own business, or filled with less productive business owners. As the neoclassical theory of labor demand stated (Hamermesh (1996)), there might be a substitution effect between capital and labor as production input.

Table 5: Model Estimates: Female Effect on Employment by Start-up Capital

	IV Estimation: Log of Employment							
	Less than $5,000	$5,000 to $9,999	$10,000 to $24,999	$25,000 to $49,999	$50,000 to $99,999	$100,000 to $249,999	$250,000 to $999,999	$ 1,000,000 or more
Female Owner	0.6529***	-0.3803	0.6662**	-0.1107	-0.3641	-0.2912	-0.6465*	-0.0290
	[0.190]	[0.353]	[0.290]	[0.299]	[0.294]	[0.196]	[0.344]	[0.282]
State Fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs	43,367	19,999	26,522	19,772	21,606	21,563	14,395	6,533
Adj R^2	0.6317	0.6524	0.6575	0.6673	0.6703	0.7025	0.7037	0.8105
F Test (df_n, df_d)	8.2274	2.2005	3.2515	0.8895	3.9437	4.3922	7.1188	0.3357
(P-value)	(0.0064)	(0.1454)	(0.0785)	(0.3510)	(0.0536)	(0.0422)	(0.0108)	(0.5654)
	IV Probit: Probability of being Employer							
Female Owner	-1.2932***	-1.9081***	-1.4397***	0.2807	-2.2136***	0.4479	0.6237	-1.9828
	[0.282]	[0.351]	[0.462]	[0.694]	[0.787]	[0.724]	[1.269]	[3.686]
State Fixed	Yes	Yes	Yes		Yes	Yes	Yes	
Industry Fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs	168,679	50,139	54,028	35,359	33,419	29,917	18,291	7,598
χ^2 Test (df)	8.1402(1)	24.2947(1)	7.9078(1)	0.3955(1)	9.9172(1)	0.7169(1)	0.4471(1)	0.2591(1)
(P-value)	(0.0043)	(0.0000)	(0.0049)	(0.5294)	(0.0016)	(0.3972)	(0.5037)	(0.6107)

Notes: Heteroskedasticity robust standard errors clustered by state are reported in square brackets. The symbols, *, **, and *** indicate respectively that the estimated coefficient is statistically significant under 10%, 5%, and 1% significance levels.

Examining female owner effect by size of start-up capital is therefore an important robustness check to confirm the positive female owner effect discussed in the previous section. Fairlie and Miranda (2017) estimate female owner effect on probability of hiring first employee by firm age, and find a consistent probability differential between female and male owners that female owners are about 10% less likely to hire first employee regardless their firm age.

The female owner effect by start-up capital size present that female owners with smaller size of startup capital are less likely to become employers. For labor demand as employers, however, start-up capital size do not seem to make effective difference in female owner effect. Table 5 reports the IV estimation of female owner effect on both log of employment and probability of being employer. By looking at the top panel of table 5, we can see that there are no consistent patterns of the female owner effect estimate along start-up capital size. The estimates in the first and third columns are positive and significant under 5%. significance level, and the estimate in the second from the right is significantly negative under 10% significance level. But the estimate with stat-up capital between $10K-25K report an acceptable p-value for the endogenous IV test under 5% significance level. On the other hand, the estimates of female owner effect on probability of being employer, reported in the bottom panel of table 5, present a consistent pattern that are significantly negative up to the start-up capital size between $50K-100K. Though, non of the significant estimates

Table 6: Model Estimates: Female Effect on Employment by Industry

	TSLS: Log of Employment					
	Female Dominant Industry			Male Dominant Industry		
	Educational Service	Health Care & Social Assistance	Retail Trade	Construction	Management	Transportation & Warehousing
Female Owner	0.3233	1.4223*	0.1296	0.1972	-0.4677	0.4737**
	[0.335]	[0.762]	[0.130]	[0.138]	[0.427]	[0.204]
State Fixed	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed	No	No	No	No	No	No
# of Obs	2,295	25,446	27,956	33,376	1,806	9,079
Adj R^2	0.6974	0.5082	0.7619	0.7826	0.6950	0.7875
χ^2 Test (df)	0.0754(1)	3.2682(1)	0.0050(1)	1.5208(1)	1.9199(1)	5.4017(1)
(P-value)	(0.7836)	(0.0000)	(0.0706)	(0.9436)	(0.1659)	(0.0201)
	IV Probit: Probability of being Employer					
Female Owner	2.5487*	-27.9199	8.0579***	1.2767***	0.3977	2.8990***
	[1.442]	[74.794]	[2.144]	[0.330]	[1.360]	[0.552]
State Fixed	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed	No	No	No	No	No	No
# of Obs	11,014	51,557	62,091	65,618	1,642	22,680
χ^2 Test (df)	8.2255(1)	9.2070(1)	157.1647(1)	24.3706(1)	0.1912(1)	45.1265(1)
(P-value)	(0.0041)	(0.0024)	(0.0000)	(0.0000)	(0.6619)	(0.0000)

Notes: Heteroskedasticity robust standard errors clustered by state are reported in square brackets. The symbols, *, **, and *** indicate respectively that the estimated coefficient is statistically significant under 10%, 5%, and 1% significance levels.

come up with acceptable endogenous IV test statistics.

Next, we estimate the female owner effect by industry. The 2007 SBO provide 2-digit north american industry classification system (NAICS) codes. We tabulate the weighted fractions of female/male owners by 2-digit NAICS code, presented in figure 2. "Educational services" and "health care and social assistance" (NAICS codes 61, 62 respectively) are the only industries that have more female owners than male. 57.80% of business in the education service industry and 56.77% for health care and social services.

Table 6 reports the female owner effect estimate by industry. We choose the top three and bottom three industries in terms of the female owner ratio. Note that "retail trade" (NAICS=44) has 48.24%, "construction" (NAILS = 23) has 7.79%, "management of companies and enterprises" (NAICS=55) has 12.29%, and "transportation and warehousing (NAICS=48) has 13.16% of female owned firms. As shown in the top and bottom panels of table 6, the estimated female effects do not seem to have any notable patterns in terms of female owned firm ratio. The estimated coefficients from TSLS for log of employment are significantly positive in the healthcare & the social assistance and the transportation & warehousing industries. From the IV probit, the estimates in the education service, retail

Table 7: Model Estimates: Female Effect on Employment by Years of Operation

	IV Estimation: Log of Employment								
	From 2007	From 2006	From 2005	From 2004	From 2003	From 2000 and 2002	From 1990 and 1999	From 1980 and 1989	From before 1980
Female Owner	1.0265	4.5780	-2.3771	0.4681	-0.5816	-0.1364	0.0464	0.1138*	0.0542
	[1.505]	[6.934]	[5.079]	[0.886]	[0.893]	[0.261]	[0.092]	[0.064]	[0.056]
State Fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs	4,911	9,840	10,635	10,816	9,010	25,021	63,095	50,532	58,161
Adj R^2	0.1985	NA	NA	0.5939	0.5815	0.6618	0.7268	0.7670	0.8111
F Test (df_n, df_d)	0.5306	4.0894	2.2301	0.1362	0.7351	0.9615	0.8735	0.1052	0.2804
(P-value)	(0.4704)	(0.0496)	(0.1428)	(0.7139)	(0.3961)	(0.3324)	(0.3553)	(0.7473)	(0.5992
	IV Probit: Probability of being Employer								
Female Owner	3.3037	6.6741	-0.5327	2.0500	-1.2775	0.7957	0.4120*	0.8211***	1.4601***
	[12.506]	[50.688]	[2.657]	[7.813]	[1.985]	[0.510]	[0.217]	[0.168]	[0.147]
State Fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs	53,197	40,697	38,049	32,842	25,263	67,913	136,556	90,321	86,813
χ^2 Test (df)	0.1866	0.1364	0.0000	0.1955	0.1539	8.3595	22.7647	78.6164	230.8919
(P-value)	(0.6657)	(0.7119)	(0.9997)	(0.6583)	(0.6948)	(0.0038)	(0.0000)	(0.0000)	(0.0000)

Notes: Heteroskedasticity robust standard errors clustered by state are reported in square brackets. The symbols, *, **, and *** indicate respectively that the estimated coefficient is statistically significant under 10%, 5%, and 1% significance levels.

trade, construction, and transportation & warehousing are significantly positive, but non of those come up with acceptable endogenous IV test statistics.

References

de Mel, S., McKenzie, D., and Woodruff, C. (2009). Are women more credit constrained? experimental evidence on gender and microenterprise returns. *American Economic Journal: Applied Economics*, pages 1–32.

Edwards, L. N. and Field-Hendrey, E. (2002). Home-based work and women's labor force decisions. *Journal of Labor Economics*, 20(1):170–200.

Fairlie, R. W. and Miranda, J. (2017). Taking the leap: The determinants of entrepreneurs hiring their first employee. *Journal of Economics & Management Strategy*, 26(1):3–34.

Hamermesh, D. S. (1996). *Labor Demand*. Princeton University Press.

Stock, J. H., Wright, J. H., and Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20(4).

Wiswall, M. and Zafar, B. (2018). Preference for the workplace, investment in human capital, and gender. *The Quarterly Journal of Economics*, 133(1):457–507.