The Gender Gap in Entrepreneurship*

J. David Brown (U.S. Census Bureau and IZA)
John S. Earle (George Mason University and IZA)
Mee Jung Kim (George Mason University and U.S. Census Bureau)
Kyung Min Lee (George Mason University and World Bank)

December 2019

Abstract
Studying detailed microdata for the U.S., we document a pronounced gender gap in business ownership and performance. Women are much less likely than men to own businesses and to participate in non-family entrepreneurial teams. Firms owned by women and gender-diverse teams distinctly lag in measured employment, exporting, and innovation. Rich data allow controls for many other characteristics of entrepreneurs, including motivations for starting a business, and of firms, including measures of access to finance. We can thus disentangle some competing explanations for the gender gap, including skills, occupational choice, entrepreneurial motivations, and financial constraints.

JEL codes: D22, J16, J24, L25, L26

1. Introduction

What role does the entrepreneur's gender play in the success of a firm? While there has been voluminous research on gender differences in the labor market generally as well as some research on self-employment, we know little about the relative success of men and women in founding businesses that create jobs and innovate, potentially contributing to employment and economic growth. We also know little about the extent of gender diversity within ownership teams, and the relationship between such diversity and firm performance. Finally, we know even less about factors that might account for an entrepreneurial gender gap, including differences in skills, choices, financial access, and motivations for entrepreneurship.

In this paper, we address these questions using newly available and unusually rich data on U.S. households and businesses. The outcome variables we analyze include employment, productivity, innovation activities, research and development (R&D), exports, and patents. Because of the high skewness in firm performance outcomes, we also examine the top end of the distributions, the high growth "gazelles" and the biggest innovators. For innovation, we measure the incidence of various activities related to product and process improvements, R&D, and patenting, and we also try to capture the degree to which the type of innovation involves exploration of truly new ideas versus exploitation of ideas from others.

In estimating differences by gender in these job creation and innovation outcomes, we are able to control for a wide variety of owner characteristics, including age, race, ethnicity, education, etc. For firms with multiple owners, we examine gender diversity in ownership teams, and the impact of diversity on this dimension for firm performance. We also distinguish gender diversity outside of from that within a married couple, and because this type of diversity is only possible with multiple owners, we pay close attention to measuring and controlling for the number of owners. In general, we are interested in describing the share and characteristics of female-owned firms and those with owner gender diversity relative to firms owned only by men, as well as in estimating the differences in firm performance.

Consistent with the small previous literature (discussed below), our own preliminary results (also discussed below) suggest that female entrepreneurs account not only for a small share of all newly created employer businesses, but also that they create fewer jobs at those businesses. We therefore consider three hypotheses that may account for this stylized fact. First, women may face discrimination in capital markets, or for other reasons have worse financial access relative to men. While we observe start-up finance in our data, and indeed we find it is on average lower for female entrepreneurs, we cannot know whether this reflects constraints faced by women, or rather choices they make about the nature of the business they are founding. Thus, a second hypothesis we investigate concerns the motivations for business ownership, and whether women are more likely to have non-pecuniary reasons such as work-family balance or flexibility, and less likely to start a business because of a new idea or for growth. Relatedly, a third hypothesis is that there are gender differences in choices made when the firm is founded that result in lower growth. Here we consider the roles played by a number of factors that may be jointly determined with firm performance outcomes, including choice of industry, legal form (including franchising), home location, and several types of founder roles in the new company. In all these cases, we are interested in measuring both how these variables differ by gender, and by diverse versus non-diverse founding teams, and in estimating how they influence or mediate the entrepreneurial gender differences in job creation and innovation performance. As we show below, a substantial

fraction of businesses (more than a quarter) are owned by married couples, which we also distinguish in constructing the categories for estimation (and which may be of independent interest).

Our analysis exploits a new data source - the Annual Survey of Entrepreneurs (ASE) – which like the Survey of Business Owners (SBO) provides detailed information on owners for a large sample of firms. The ASE allows us to measure several sets of variables not available in the SBO: motivations for ownership, innovation activities, research and development, and patenting. We also link the ASE data to the Census Bureau Business Register (BR) and Longitudinal Database (LBD) to measure firm age and revenue per worker, and we link to patent files housed at the Census Bureau for a count of the number of patents per firm.

Our research in this paper is related to the broad topic of gender and the labor market and to specific research on gender differences in self-employment. Devine (1994a), Manser and Picot (1999), and Hipple (2010) provide basic summaries documenting lower self-employment rates among women in the U.S. Devine (1994b) and Budig (2006) analyze relative earnings of female self-employed and employees, including the gender gap in employee earnings as a possible motivation for female self-employment. Hundley (2000), Lombard (2001), and Wellington (2006) attempt to discern the influence of motivations for self-employment by studying marital status and other family characteristics, the latter focusing on married women. Gurley-Calvez, Biehl and Harper (2009) compare time use of self-employment men and women with a similar purpose.

We also contribute to an emerging literature on entrepreneurship and job creation using broad, representative samples of firms. Jarmin and Krizan (2010) find that women-owned businesses have lower average employment growth rates in the 2002 SBO linked to the LBD. Using the same data, Jarmin, Krizan and Luque (2014) also report that women-owned business underperform on employment growth during the Great Recession. Using the 1992 Characteristics of Business Owners (CBO), the predecessor survey to the SBO, Fairlie and Robb (2009) report women have a lower probability of hiring employees, among other measures of business success. Analyzing an entry cohort from the 2007 SBO linked to the LBD, Brown, Earle, Kim and Lee (2019) find female-owned firms are 50% less likely to be in the top 5% of the employment distribution. Blanchard, Zhao, and Yinger (2008) and Coleman and Robb (2009) study financial constraints of female entrepreneurs. As an example of much of the literature using small sample surveys, Kalleberg and Leicht (1991) find small, statistically insignificant disadvantages of women in survival and earnings growth by gender of owner in a survey they conducted of 411 firms, 99 of them owned by women. These studies examine cross-sections, not distinguishing firm size or growth by firm age. There appears to be little analysis of gender diversity in ownership teams using large samples.[1] Finally, our paper is the first to use nationally representative data to examine innovation and productivity outcomes of entrepreneurship by gender and gender diversity.

The following section describes the data in more detail. Section 3 contains methods and Section 4 contains results. Section 5 concludes the paper.

2. Data

We use data from the Census Bureau's 2014 Annual Survey of Entrepreneurs (ASE). The ASE is an annual survey that supplements the Survey of Business Owners (SBO), conducted every

---

[1] But see Hoogendorn and van Praag (2013). Ruef, Aldrich and Carter (2003) study team size. Lazear (2004, 2005) argues that diverse skills are valuable for entrepreneurs which may be substituted by a diverse team. A related but distinct literature concerns diversity in management teams, such as Adams and Ferreira (2009) and Flabbi, Macis, Moro, and Schivardi (2019).

five years. The survey contains a nationally representative random sample of non-farm businesses with at least one paid employee and receipts of $1,000 or more. Using the Census Business Register (BR) as the sampling frame, the sample is stratified by the 50 most populous Metropolitan Statistical Areas (MSAs), state, and the firm's number of years in business.[2] In each stratum, large firms are selected with certainty based on the volume of sales, payroll, or number of paid employees. The ASE provides detailed demographic characteristics on business owners (up to four) and their motivations to start a business, as well as economic characteristics of their firms. It also includes various information related with business performance such as job creation, innovation measures, and R&D activities. The initial 2014 ASE sample is about 290,000 employer firms, and the response rate is about 74 percent (Foster and Norman 2016).

The unit of analysis in our study is business owners. We exclude businesses where no individual owns at least 10 percent of the equity since owner information is not provided for such businesses. We also drop owners who choose the same answers for every motivation question (all very important, all somewhat important, or all not important), considering those answering patterns may not reflect the true intensity for each question. Finally, we restrict our sample to firm-owner observations that have non-missing values for any of the variables used in the regressions. Our final sample consists of about 179,000 owners of 119,000 employer businesses. (our export sample consists of about 140,000 owners of 93,000 employer businesses). We weight owner level observations by their equity share to make each observation representing its share of the firm, and then we weight each firm by the survey weights to make the sample representative for the U.S. economy.

The main explanatory variable of interest is female ownership, but we need to take into account cases with multiple owners, including those with gender diversity and those where there is a family relationship (most commonly, a married couple). To measure the ownership structure by gender, we constructed six ownership variables, as shown in Figure 1: single female, multiple females, family, unrelated owners, single male, and multiple males.
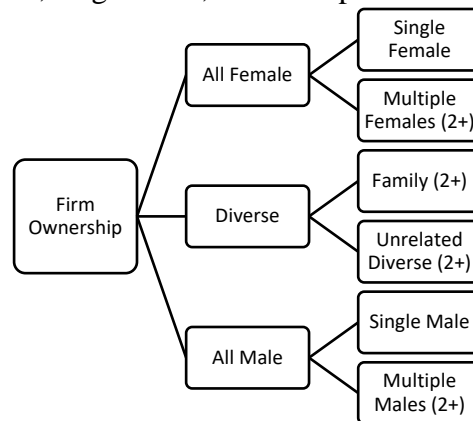

Figure 1. Definition of Ownership Categories

Single female and single male ownership variables represent one owner and other ownership variables reflect two or more owners. Multiple female owners indicate whether two or more firm owners are all female owners, whereas multiple male owners capture whether owners in the firm with two or more owners are all male owners. It is important to aware that gender diversity only applies to the firms with two owners or more with different genders. The family variable indicates

---

[2] See Foster and Norman (2016) for further details about the ASE.

whether there is gender diversity among owners who are relatives.[3] Unrelated owners represent the mix of multiple male and female owners without family relationships. In the analysis, we use the single male owner as a base category and always include other five ownership variables together.

Sets of control variables, available for all owners and firms in the sample, include owner demographics, human capital, motivations, start-up capital, industries, and other choices about the owner's participation in the business. Owner demographics comprise age and race/ethnicity. Age is a categorical variable for six age groups (less than 25, 25-34, 35-44, 45-54, 55-65, and 65 or over). For race/ethnicity, we first construct a dummy variable for Hispanic and then create non-Hispanic racial groups including White, African American, Asian, and other minorities (Native Hawaiian, Guamanian or Chamorro, Samoan, other Pacific Islander, and some other race). Human capital variables are veteran status, prior business experience, and a categorical variable for educational attainment prior to owning the business.

The ASE has nine separate questions on motivations for owning the business, including 1) "Best avenue for my ideas/goods/services" (*Ideas*); 2) "Opportunity for greater income/wanted to build wealth" (*Income*); 3) "Couldn't find a job/unable to find employment" (*No Job*); 4) "Wanted to be my own boss" (*Own Boss*); 5) "Working for someone else didn't appeal to me" (*Work for Self*); 6) "Always wanted to start my own business" (*Always Wanted*); 7) "An entrepreneurial friend or family member was a role model" (*Role Model*); 8) "Flexible hours" (*Flexible Hours*); 9) "Balance work and family" (*Balance Family*). We construct a dummy variable for each motivation, which is equal to 1 if somewhat or very important and 0 otherwise.

The ASE also provides information on the amount and source of start-up finance as well as industries (4-digit NAICS). The amount of start-up capital is a categorical variable from less than $5,000 to $3 million or more. The start-up funding sources include a bank or financial institution, venture capitalist(s), government-guaranteed business loans, or grants. Furthermore, the ASE has detailed information on access to finance. It asks whether the owner chose not to apply when in need of additional financing, as well as the reasons for not applying, such as discouragement from expected nonapproval, or not wanting to accrue debt. This allows us to identify discouraged borrowers as well as their corresponding reasons.

Other choice variables include job function in business, primary income, hours worked, home-based, full-time operation, and franchise. Job functions are a categorical variable for the owner's main role in the business including manager, good/service provider, financial controller, and none of these roles. Primary income is a dummy variable indicating whether this business is the owner's primary income source. Hours worked is a categorical variable for average weekly hours the owner spends managing or working in the business. Full-time operation is a dummy variable to specify whether the business operates full-time in 2014 vs. less than 40 hours per week, less than 12 months, seasonal business, or operates occasionally. Home-based is a dummy variable indicating whether the business operates primarily from somebody's home in 2014.

3. Methods

We estimate a series of regression models to study the gender gap in entrepreneurship. We use a linear probability model for binary innovation outcomes, linear regression for continuous performance variables, and Poisson regression for the number of patents. The regressions are estimated via weighted least squares with standard errors clustered by firm. The firm-owner

---

[3] Brannon, Wiklund and Haynie (2013) study family relationships in entrepreneurial teams.

observations are weighted by ownership shares, and then firms are weighted by the ASE weight to be representative to the U.S. economy. Our baseline model is specified as follows.

$$Y_{ij} = \alpha + \beta F_{ij} + \tau' T_j + f(Age_j) + Z_j + u_{ij}, \tag{1}$$

where Y_{ij} is an outcome variable (employment, productivity, innovation measure, etc.) for firm j owned by owner i. The main variable of interest F_{ij} is an indicator if owner i is female. Taking into account team size and diversity, T_j is a vector of team types for firm j, including dummy variables for multiple female owners (2+), multiple male owners (2+), and family (2+), and unrelated multi-owners (2+). The reference category is a single male owner. Since businesses are at different ages, and firm performance measures are related with firm age, we control for a quadratic function of firm age, $f(Age_j)$.

The dependent variables Y_{ij} include the number of employees in 2014 and a dummy variable whether the firm's number of employees is in the top 5 percent of the LBD employment distribution in 2014. Regarding innovation, the dependent variables include dummies for whether the firm engages in any innovation activities, for any R&D activities, and for several types of product and process innovations; they also include the number of patents. Additional dependent variables are an indicator for whether the firm has a sales of goods and/or services of exports outside the U.S. in 2014 and an indicator for whether the owner's business aspiration is to be larger in terms of sales or profits in five years.

The coefficient of the female owner indicator (β) captures the differences of outcomes between single female-owned and single male-owned firms. The estimates describe the differences in firm-level outcomes between single female-owned and single male-owned firms. Just as in an analysis of gender differences in wages, for example, there is no issue of causality: we do not interpret the results as the impact of turning a random male-owned into a female-owned (just as the interpretation placed on a female coefficient is not the impact of changing a male into a female).

After describing the raw gaps using equation (1), we estimate how much these gaps are accounted for by observable differences in demographic and human capital characteristics. For this purpose, we estimate the following specification:

$$Y_{ij} = \alpha + \beta F_{ij} + \tau' T_j + f(Age_j) + Z_{ij} + X_{ij}\lambda + u_{ij} \tag{2}$$

where X_{ij} is a vector of characteristics of i of firm j. The vector includes demographic variables (gender, age, and race/ethnicity) and proxies for human capital (education, veteran, and prior business). The β estimated from equation (2) captures the magnitude of the gender gap in business outcomes adjusted for owner characteristics. Even after controlling for exogenous owner characteristics, our gender gap estimates may reflect differences in other observables that are less clearly exogenous and indeed may be jointly determined with a range of outcomes. Examples of such observables include motivations, start-up capital, industries, and other choices.

Next, we add several sets of variables that may help shed light on ownership decisions but that are also potentially endogenous to the decision. The first set is motivations for ownership. Many small businesses are started with non-pecuniary motives and with no intention to grow or innovate (Hurst and Pugsley 2011). The literature suggests that female entrepreneurs with a motivation to balance work and life are less likely to succeed (e.g. Rey-Martí et al. 2015). For this purpose, we estimate the following specification:

$$Y_{ij} = \alpha + \beta F_{ij} + \tau' T_j + f(Age_j) + Z_{ij} + X_{ij}\lambda + M_{ij} + u_{ij} \tag{3}$$

where M_{ij} is the set of motivation variables. As described in the data section, the survey asks about reasons of owning a business. Nine motivation variables include idea, income, no job, own boss, work for self, always wanted, role model, flexible hours, and balance work and life. The survey respondents indicate whether each motivation is not important, somewhat important, or very

important. In our specifications, we include dummies for somewhat and very important for each motivation.

We then add controls for the amount of start-up capital for businesses. The importance of access to finance for business start-ups is well documented in the literature (e.g. Evans and Jovanovic 1989; Evans and Leighton 1989). The higher start-up finance among female owners may account for the differences in business outcomes between female- and male-owned firms. The regression model is specified as follows.

$$Y_{ij} = \alpha + \beta F_{ij} + \tau' T_j + f(Age_j) + Z_{ij} + X_{ij}\lambda + M_{ij} + K_j + u_{ij} \qquad (4)$$

where K_j is the set of vectors of the amount of start-up finance. The amount of start-up finance available in the ASE is a categorical variable with options from less than \$5,000 to \$3 million or more as well as an option for "don't need." We use the category for less than \$5,000 as our reference category and include dummy variables for all the other categories.

We also control for heterogeneity of industry choices. Female owned businesses are concentrated in retail and service industries where the average size of businesses are smaller than those in other industries (Anna et al. 2000). Different selection processes into specific industries may explain the part of the gender gap in entrepreneurship. We estimate the following specification.

$$Y_{ij} = \alpha + \beta F_{ij} + \tau' T_j + f(Age_j) + Z_{ij} + X_{ij}\lambda + M_{ij} + K_j + S_j + u_{ij} \qquad (5)$$

where S_j represents the industry fixed effects. We use the 4-digit NAICS industry codes, which allows us to compare female- and male-owned firms within the same industry.

Finally, we control for heterogeneity in several types of owner choices, which may explain the remaining gender gap in outcomes.

$$Y_{ij} = \alpha + \beta F_{ij} + \tau' T_j + f(Age_j) + Z_{ij} + X_{ij}\lambda + M_{ij} + K_j + S_j + C_{ij} + u_{ij} \qquad (6)$$

where C_{ij} include the set of choice variables including job function in business, primary income, hours worked, home-based, full-time operation, and franchise. The empirical exercise from equation (1) to (6) allows us to understand how much gender gap in entrepreneurship is explained by the sets of the observable characteristics. Furthermore, our estimate of the gender gap in equation (6) shows the remaining gender gap after accounting for an exhaustive set of observable characteristics and related choices.

4. Results
4.1. Descriptive Statistics

Table 1 provides descriptive statistics for owner characteristics. The first column includes estimates for all sample and the remaining three columns divide the sample into three sub-samples including all female-, diverse team-, and no female-owned firms. The descriptive statistics show that the percentages of owners in single female- and male-owned firms are about 13 and 43 percent, respectively. The female owners having multiple female-owned businesses cover the very small shares, which are only about 1.7 percent. About 28 percent of owners have the firms with diverse owners, among which the share of family-owned businesses is almost 5 times larger than firms with unrelated owners.

The age categories show that the share of owners increase by age 55-65 and then decrease, which suggest an inverse u shape relationship between age and the propensity of owning a business. The female-owned businesses tend to have young owners relative to diverse team- or no female-owned businesses. The major race/ethnicity group is White, which covers about 84 percent. The second largest race/ethnic group is Asian (about 8.4 percent) and the third group is Hispanic (about 5.2 percent). The percentage of black owners is only about 1.5. The percentage of non-White

owners is larger in firms with all female owners compared to those with diverse- or no female-owners.

Concerning education, a slightly higher share of female owners have graduate degrees. On the other hand, the share of owners with prior business experience or veteran experience much smaller in the female-owned businesses.

Table 2 presents the descriptive statistics of motivations to own businesses. Females have a much higher propensity to own the business because they wanted to have either flexible hours or balance work and family, or because could not find a job. Relevant for innovation, females report slightly higher propensity to own a business as the best avenue for their ideas, goods, or services.

In Table 3, female-owned businesses tend to use less capital to start the business, suggesting that financial market discrimination may be a problem for female owners. In this regard, female owners have a higher propensity of having startup funding from grants, whereas diverse and male owners use more startup funding from bank or financial institution or venture capitals. Female-owned firms have a higher propensity to apply additional funding from grants. However, female-owned firms are more likely to be discouraged to apply for additional funding because they don't think that would be approved by lenders.

We show the relative propensity of female ownership by broad sectors in Table 4. As some 2-digit NAICS sectors have small numbers of firms in the dataset, we combine agriculture and mining sector as the primary sector, we add utilities to the manufacturing sector, and the administrative sector includes management of companies. Female-owned businesses are much more prevalent in business services and health care, and they are also somewhat more represented in the retail sector.

In Table 5, descriptive statistics for other choices are presented. Female-owners have a higher propensity of holding different roles of manager, service provider, and financial controller. They are more likely to work full-time, having lower propensity to work more than 40 hours compared to male owners. Additionally, female owners have a higher propensity to choose the business as a primary income and home-based relative to male-owned firms. Perhaps, related to the non-pecuniary motivation of owning the businesses, female-owned firms have lower aspiration to be larger for their businesses compared to male-owned or diverse ownership firms.

Table 6 shows the descriptive statistics for firm performance measures including innovation, employment, aspiration to grow, and export. Our outcome variables include detailed innovation, research and development (R&D), and number of patents measures. The ASE asks whether the business conducted twelve different product or process innovation activities in the last three years (2012-2014). We create a binary variable for innovation to indicate whether a firm conducted any product or process innovation in the last three years. We also create the count of innovations by aggregating the number of product and process innovation activities. We create binary indicator variables for each type of product and process innovation activities. Particularly, product innovations include whether the firm sold a new good or service that no other business has ever offered before and whether the firm sold a new good or service that this business has never offered before. Process innovations include whether the firm applied a new way of purchasing, accounting, computing, maintenance, inventory control, or other support activity and whether the firm decreased production costs by improving the materials, software, or other components. About 57 percent of firms conducted at least one innovation, and the unconditional mean of the number of innovation types is 2 of the 12 in our sample. Female-owned firms have a lower propensity to

conduct innovation activities across different kinds of product and process innovation compared to male-owned firms.

For R&D activities, the ASE asks business owners to answer whether their business did seven different R&D activities in 2014. We create an indicator for whether the business conducted any of these R&D activities. Seven R&D activities include 1) conducted work that might lead to a patent; 2) developed and tested prototypes that were derived from scientific research or technical findings; 3) applied scientific or technical knowledge in a way that has never been done before; 4) produced findings that could be published in academic journals or presented at scientific conferences; 5) created new scientific research or technical solutions that can be generalized to other situations; 6) conducted work to discover previously unknown scientific facts, structures, or relationships; and 7) conducted work to extend the understanding of scientific facts, relationships or principles in a way that could be useful to others. The last set of outcome variables regard patents. Using the Business Dynamic Statistics- Patenting Firm (BDS-PF) crosswalk, we link this to the LBD to have number of patents that firms owned by 2014. Looking at Table 6, female-owned firms are less likely to conduct any R&D activities and hold patents compared to male-owned firms or firms with diverse ownership.

Figure 2 shows graphical tabulations of the female share in the population and several types of employment and self-employment. Data come from the outgoing rotation groups of the Current Population Survey (CPS) from 2014-2018 (and for comparison where possible, from 1994). While the share in total employment is only a little shy of .5, the self-employment share is little more than one-third. Female self-employed are more likely to be unincorporated, and therefore the incorporated self-employment share is under .3. These shares show only slight increases from 1994 to 2014-2018. The employer share, available only since 2014, is less than a quarter and as we consider larger thresholds for numbers of employees, the female share shrinks steadily, falling to about 18 percent for self-employed with more than 50 employees. To conclude, women are less likely to be self-employed, still less likely to own incorporated businesses, much less likely to hire employees, and still much less likely to hire employees in large numbers.

4.2. Regression Results

Next we turn to results on job creation in the ASE firm-level data on employers. Figure 3a shows that firms with any female ownership have higher employment on average than all male-owned firms. This effect appears to be driven by the fact that most firms with any female ownership have multiple owners, and firms with multiple owners tend to be larger – once controlling for teams, the female effect becomes negative. The negative female effect controlling for teams is not due to differences in education or other demographics. Differences in motivation explain over half the gap. As discussed above, women owners are more often motivated by flexible hours and balancing work and family, traits associated with smaller firms. Once controlling for start-up finance, the gender gap becomes insignificant, reflecting less start-up finance in female-owned businesses on average. The gender gap becomes significant again when controlling industry, consistent with female-owned businesses being in sectors with larger average firm size (e.g., the male owner-dominated construction industry has low average firm size). When using the top 5 percent of employment dummy as the dependent variable, the any female coefficients are always negative, and the gaps are larger. All-male multi-owner firms have the largest average employment (Figure 3b). This effect is partly explained by having greater start-up finance. Multiple-owner businesses with all female owners and unrelated mixed gender owners have similar employment effects. Mixed gender family businesses tend to be smaller than other multi-owner firms, though

they are still larger than male single-owner firms. The gaps between the all-male multi-owner firm coefficients and the other multi-owner categories are much larger in the regressions with the largest 5 percent of employment dummy (Figure 3c). The other multi-owner categories are generally insignificantly different from single male owners.

 Firms with any female ownership are more likely to engage in some form of innovation, as shown in Figure 4a. Once controlling for multiple owners, the any female coefficient drops sharply and becomes insignificant. As with job creation, multi-owner firms have a higher propensity to innovate. The any female effect further declines when controlling for demographics, but it becomes significant again once controlling for motivations (female owners are less likely to start the business to carry out new ideas), and it rises further with start-up finance (female-owned firms tend to have less). The different types of multi-owner firms have similar propensities to innovate, all greater than single male-owned firms (Figure 4b). This pattern varies little as controls are added. Unlike with the innovation dummy, the any female effects for innovation count regressions with controls are always insignificant (Figure 5a). Male multi-owner effects are the largest in all specifications, and female multi-owners tend to be the smallest among multi-owner categories (Figure 5b).

Figures 6a-11b show regression results with different types of innovations, R&D activity, and the number of patents. The gender effects vary considerably across innovation measures. Any female effects are positive for selling a good/service that is new to this firm in all specifications, and all female multi-owners and mixed gender family-owned firms have stronger effects than other types of multi-owner firms. For the process innovation of applying a new way of purchasing, accounting, computing, maintenance, inventory control, or other support activity, any female effects are also positive in all specifications, and all female multi-owners again have the strongest effects among multi-owner firms. The female coefficients are insignificant for selling a good or service no firm had offered before and for the number of patents owned, while all female multi-owner coefficients are the lowest among multi-owner categories (all male multi-owner firm coefficients are the largest). For the process innovation of decreased production costs and for a dummy for any R&D activity, the any female effects are negative, and all female multi-owner firms have the lowest propensities among multi-owner firms. All male multi-owner firms have the strongest effects for both of these dependent variables.

5. Conclusion

This paper provides new evidence on the gender gap in entrepreneurship. Although there is a sizable literature on gender differences in self-employment, and a smaller one on differences in firm performance between businesses owned by women versus those owned by men, we are able to study many more measures and dimensions of entrepreneurial success than were available to previous researchers. The outcome variables we analyze include employment, productivity, innovation activities, research and development, exports, and patents. Because of the high skewness in firm performance outcomes, we also examine the top end of the distributions, the high growth "gazelles" and the biggest innovators. For innovation, we measure the incidence of various product and process changes, R&D, and patenting, and we also try to capture the degree to which the type of innovation involves exploration of truly new ideas versus exploitation of ideas from others.

In estimating differences by gender in these job creation and innovation outcomes, we are able to control for a wide variety of owner characteristics, including age, race, ethnicity, education, etc. For firms with multiple owners, we examine gender diversity in ownership teams, and the

impact of diversity on this dimension for firm performance. We also distinguish gender diversity outside of marriage from that within a married couple, and because this type of diversity is only possible with multiple owners, we pay close attention to measuring and controlling for the number of owners. In general, we are interested in describing the share and characteristics of female-owned firms and those with owner gender diversity relative to firms owned only by men, as well as in estimating the differences in firm performance.

We find that female entrepreneurs not only have smaller probabilities of starting businesses, but they also are less likely to create employer businesses; among employer businesses, they create fewer jobs at those businesses. Female-owned businesses also tend to be less innovative, engage less in research and development, patent less often, and export less frequently. We provide further evidence on three hypotheses that may account for these systematic gender gaps. First, we examine differences in financial access, perhaps reflecting discrimination in capital markets. Second, we analyze unusual information on the reported motivations for business ownership, including whether women are more likely to have non-pecuniary reasons such as work-family balance or flexibility, or less likely to start a business as an avenue for a new idea or for growth. Finally, we study gender differences in owner choices that may result in lower growth, including the roles played by a number of factors that may be jointly determined with firm performance outcomes, including choice of industry, legal form (including franchising), home location, and several types of founder roles in the new company. We should caution that our results on these hypotheses are preliminary; we look forward to developing them further.

References

Adams, Renée B., and Daniel Ferreira. 2009. "Women in the Boardroom and their Impact on Governance and Performance." *Journal of Financial Economics* 94 (2): 291-309.

Anna, Alexandra L., Gaylen N. Chandler, Erik Jansen, and Neal P. Mero. 2000. "Women Business Owners in Traditional and Non-traditional Industries." *Journal of Business Venturing* 15 (3): 279-303.

Blanchard, Lloyd, Bo Zhao, and John Yinger. 2008. "Do Lenders Discriminate against Minority and Woman Entrepreneurs?" *Journal of Urban Economics* 63 (2): 467–97.

Brannon, David L., Johan Wiklund, and J. Michael Haynie. 2013. "The Varying Effects of Family Relationships in Entrepreneurial Teams." *Entrepreneurship Theory and Practice* 37 (1): 107-132.

Brown, J. David, John S. Earle, Mee Jung Kim, and Kyung Min Lee. 2019. "Start-ups, Job creation, and Founder Characteristics." *Industrial and Corporate Change* 28 (6): 1637-1672.

Budig, Michelle J. 2006 "Gender, Self-employment, and Earnings: The Interlocking Structures of Family and Professional Status." *Gender & Society* 20 (6): 725-753.

Coleman, Susan, and Alicia Robb. 2009. "A Comparison of New Firm Financing by Gender: Evidence from the Kauffman Firm Survey Data." *Small Business Economics* 33 (4): 397.

Devine, Theresa J. 1994a. "Characteristics of Self-employed Women in the United States." *Monthly Labor Review* 117: 20.

Devine, Theresa J. 1994b. "Changes in Wage-and-Salary Returns to Skill and the Recent Rise in Female Self-employment." *The American Economic Review* 84 (2): 108-113.

Evans, David S., and Boyan Jovanovic. 1989. "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints." *Journal of Political Economy* 97 (4): 808-827.

Evans, David S., and Linda S. Leighton. 1989. "Some Empirical Aspects of Entrepreneurship." *American Economic Review* 79 (3): 519-535.

Fairlie, Robert W., and Alicia M. Robb. 2009. "Gender Differences in Business Performance: Evidence from the Characteristics of Business Owners Survey." *Small Business Economics* 33 (4): 375-395.

Flabbi, Luca, Mario Macis, Andrea Moro, and Fabiano Schivardi. 2019. "Do Female Executives Make a Difference? The Impact of Female Leadership on Gender Gaps and Firm Performance." *The Economic Journal* 129 (622): 2390-2423.

Foster, Lucia, and Patrice Norman. 2016. "The Annual Survey of Entrepreneurs: An Introduction." CES Discussion Paper 15-40R.

Gurley-Calvez, Tami, Amelia Biehl, and Katherine Harper. 2009. "Time-use Patterns and Women Entrepreneurs." *American Economic Review* 99 (2): 139-144.

Hipple, Steven F. 2010. "Self-employment in the United States." *Monthly Labor Review* 133 (9): 17-32.

Hoogendoorn, Sander, Hessel Oosterbeek, and Mirjam Van Praag. 2013. "The Impact of Gender Diversity on the Performance of Business Teams: Evidence from a Field Experiment." *Management Science* 59 (7): 1514-1528.

Hundley, Greg. 2000. "Male/female Earnings Differences in Self-employment: The Effects of Marriage, Children, and the Household Division of Labor." *ILR Review* 54 (1): 95-114.

Hurst, Erik, and Benjamin W. Pugsley. 2011. "What Do Small Businesses Do?" *Brookings Papers on Economic Activity*: 73.

Jarmin, Ron, and C.J. Krizan. 2010. "Past Experience and Future Success: New Evidence on Owner Characteristics and Firm Performance." CES Discussion Paper 10-24.

Jarmin, Ron, C.J. Krizan, and Adela Luque. 2014. "Owner Characteristics and Firm Performance During the Great Recession." CES Discussion Paper 14-36.

Kalleberg, Arne L., and Kevin T. Leicht. 1991. "Gender and Organizational Performance: Determinants of Small Business Survival and Success," *Academy of Management Journal* 34 (1): 136–161.

Lazear, Edward P. 2004. "Balanced Skills and Entrepreneurship." *American Economic Review* 94: 208-211.

Lazear Edward P. 2005. "Entrepreneurship." *Journal of Labor Economics* 23: 649-680.

Lombard, Karen V. 2001. "Female Self-employment and Demand for Flexible, Nonstandard Work Schedules." *Economic Inquiry* 39 (2): 214-237.

Manser, Marilyn E., and Garnett Picot. 1999. "The Role of Self-employment in US and Canadian Job Growth." *Monthly Labor Review* 122: 10.

Rey-Martí, Andrea, Ana Tur Porcar, and Alicia Mas-Tur. 2015. "Linking Female Entrepreneurs' Motivation to Business Survival." *Journal of Business Research* 68 (4): 810-814.

Ruef, Marin, Howard E. Aldrich, and Nancy M. Carter. 2003. "The Structure of Founding Teams: Homophily, Strong Ties, and Isolation Among U.S. Entrepreneurs." *American Sociological Review* 68 (2): 195-222.

Wellington, Alison J. 2006. "Self-employment: The New Solution for Balancing Family and Career?" *Labour Economics* 13 (3): 357-386.
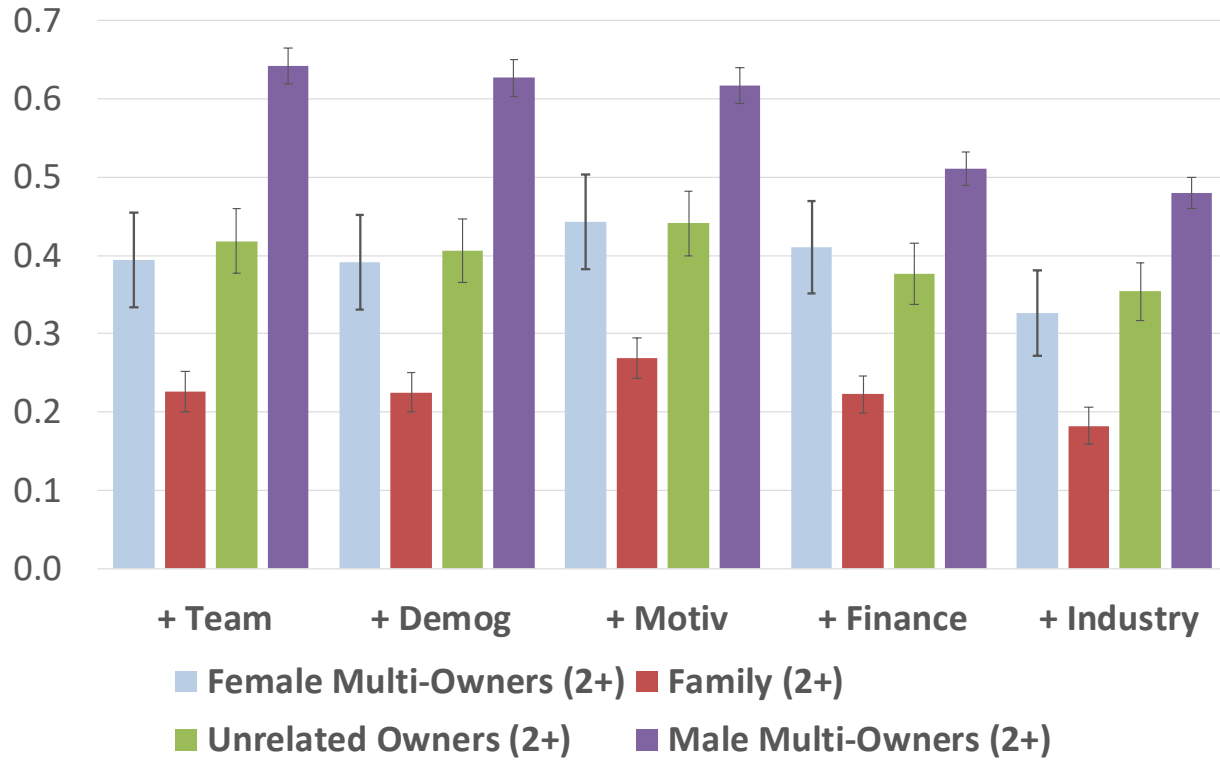
Figure 2. Share of Female in CPS 2014-2018
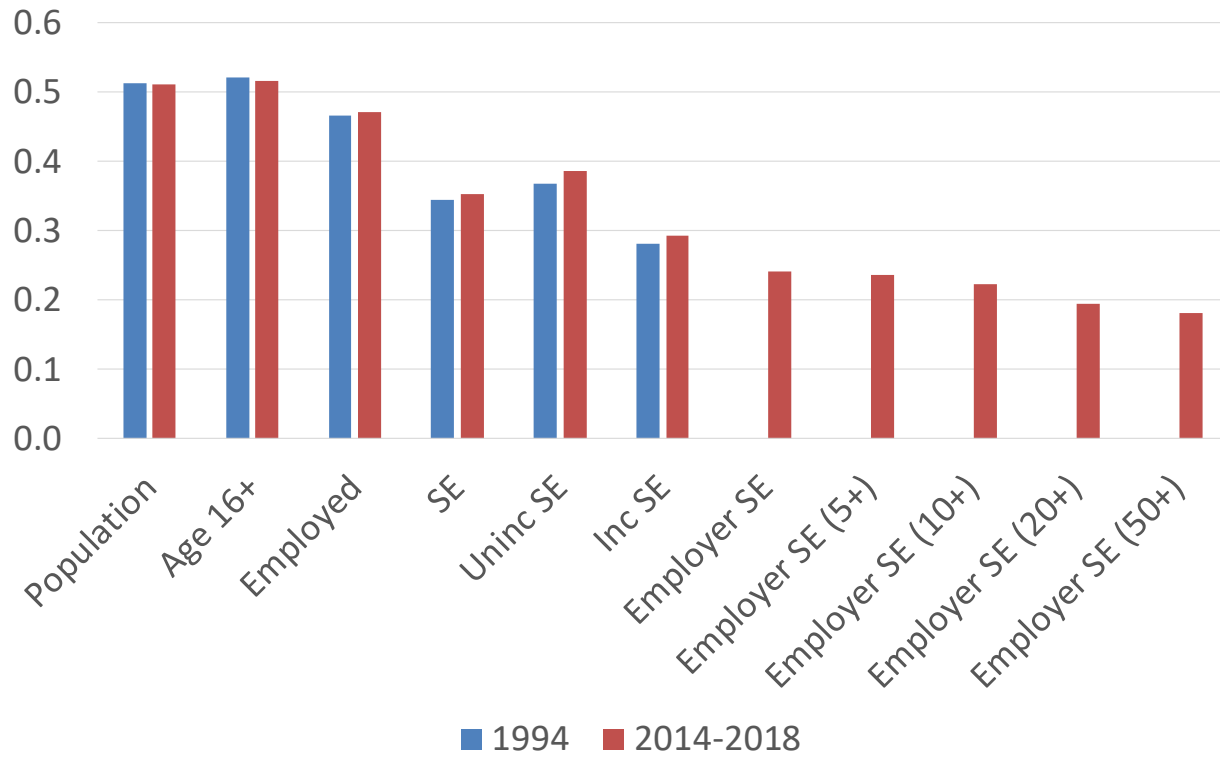
Source: CPS PUMS Merged Outgoing Rotation Group data. Data on employer status and number of employees are unavailable prior to 2014.

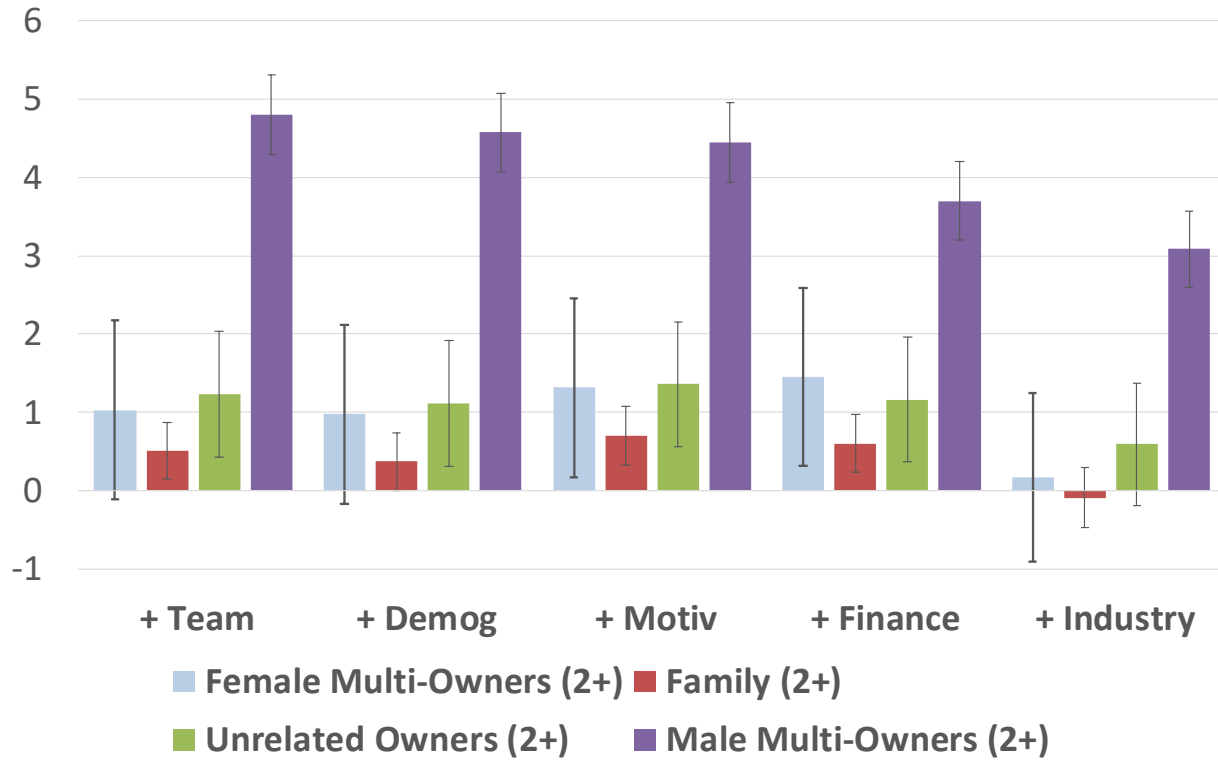**Figure 3a. Job Creation (Any Female vs. All Male)**
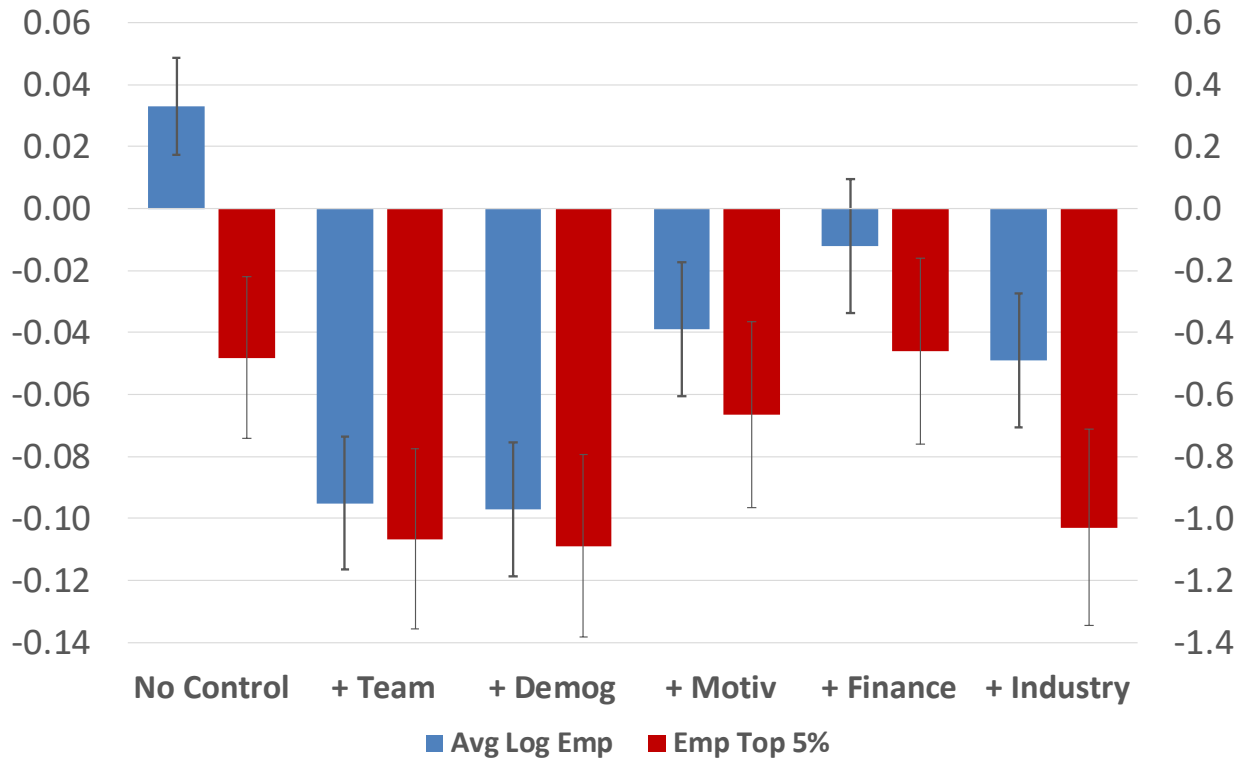
Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430

Figure 3b. Job Creation (Disaggregated Categories vs. Single Male Owner, Log Employment)
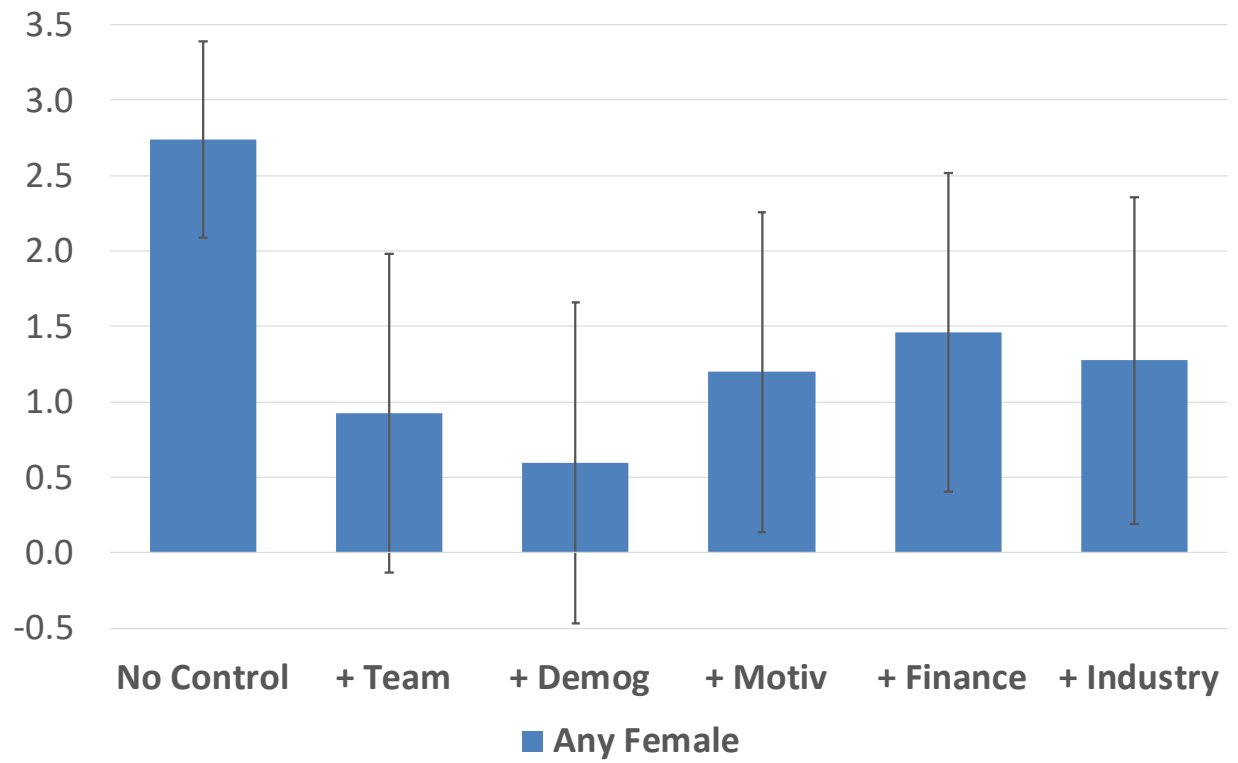

Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430

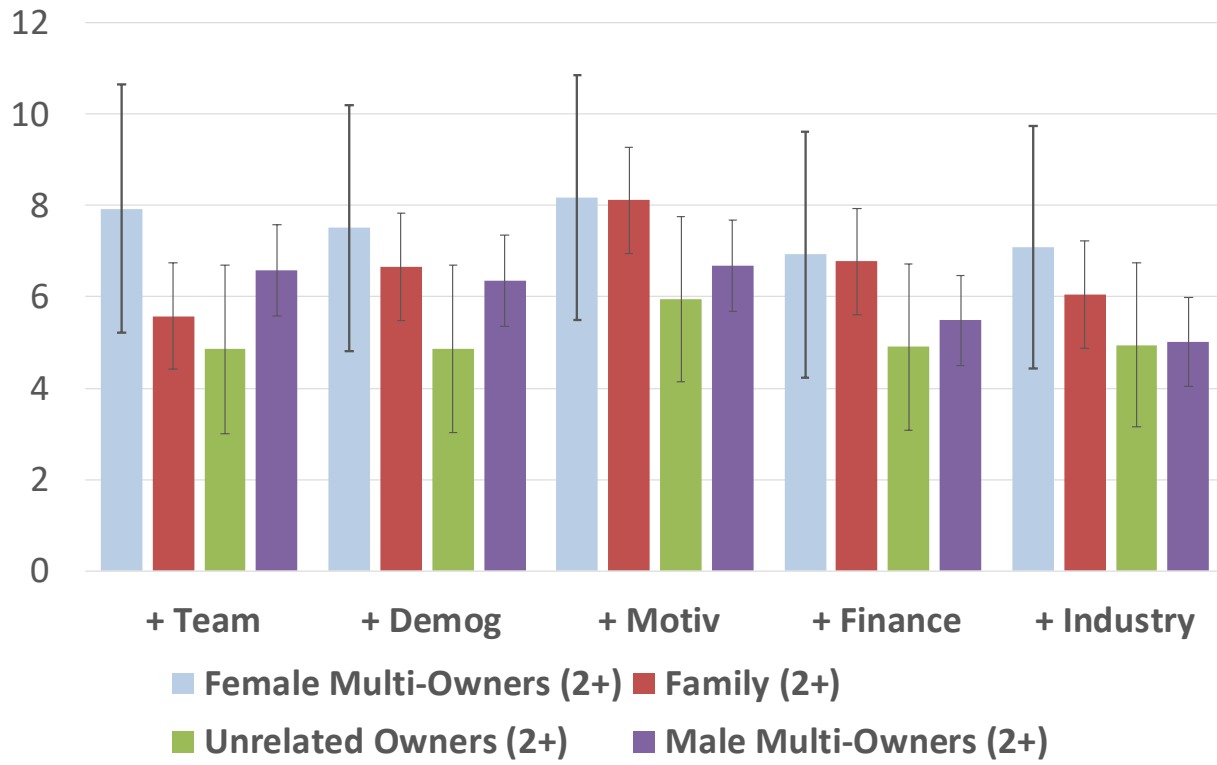**Figure 3c. Job Creation (Disaggregated Categories vs. Single Male Owner, Largest 5%)**

Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430

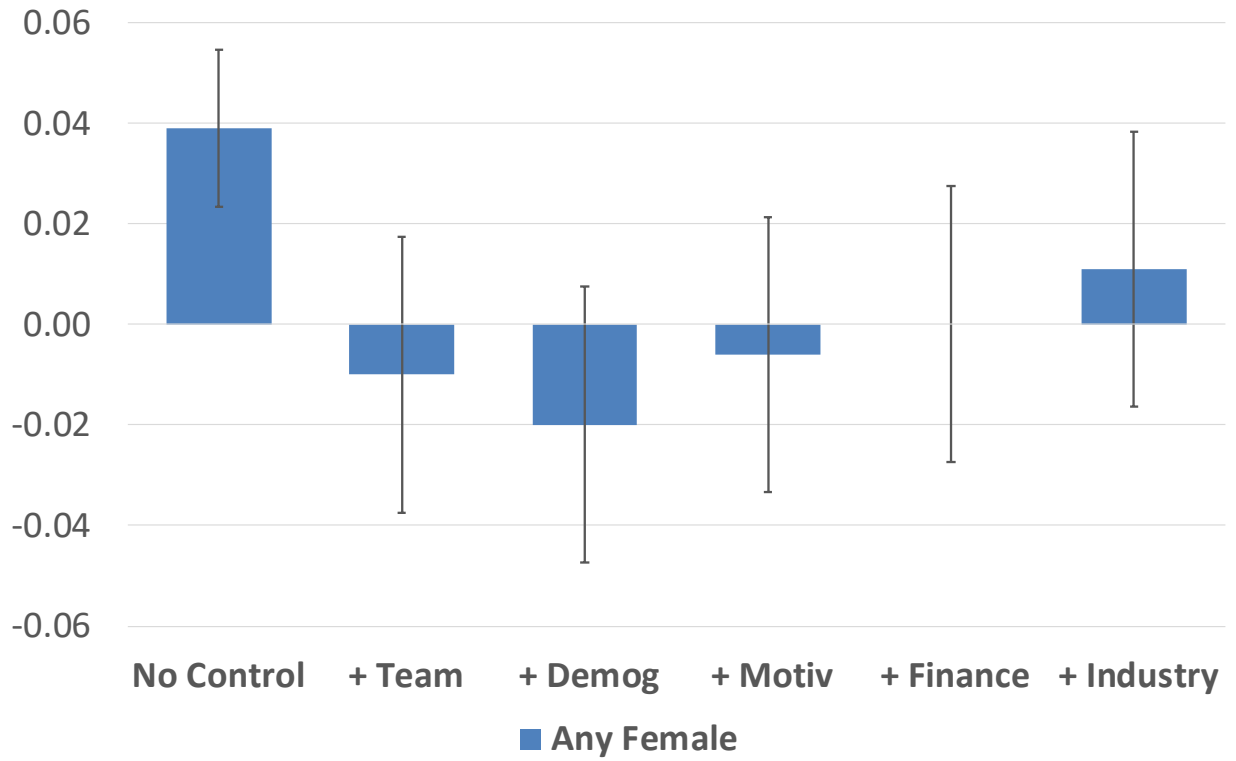**Figure 4a. Innovation Dummy (Any Female vs. All Male)**

Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430

Figure 4b. Innovation Dummy (Disaggregated Categories vs. Single Male Owner)
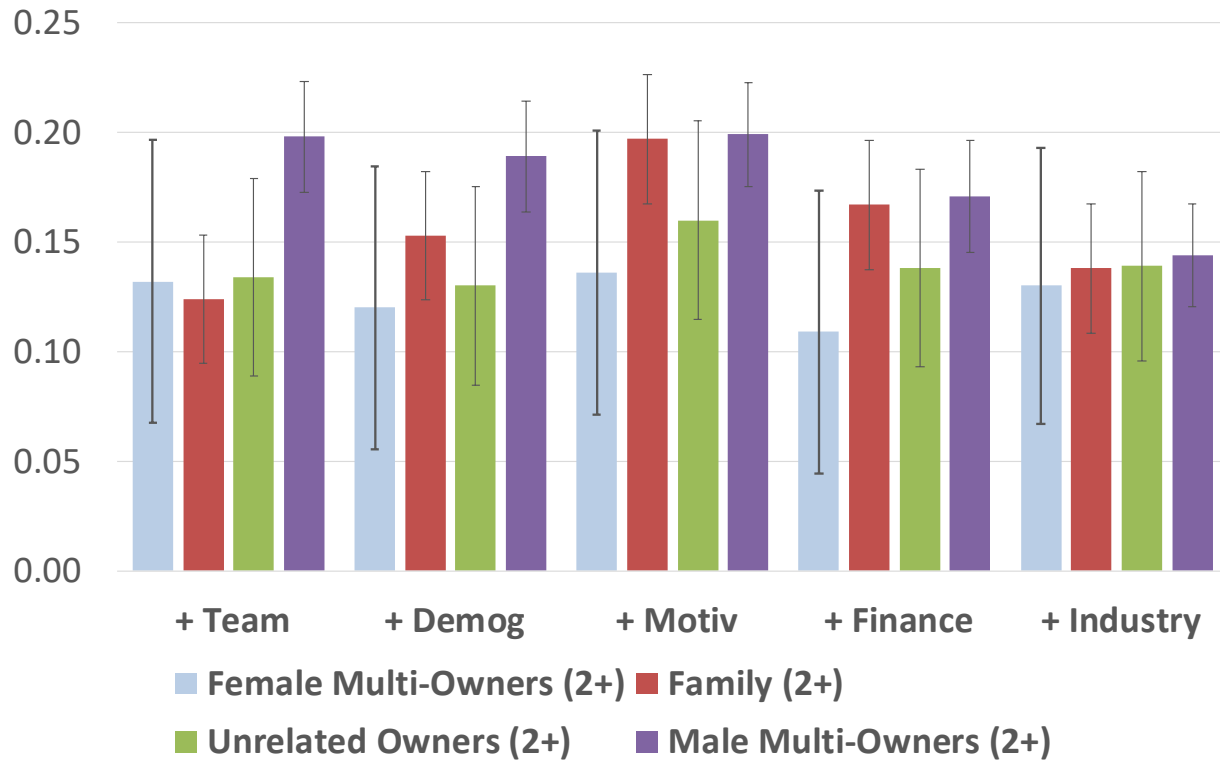

Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430

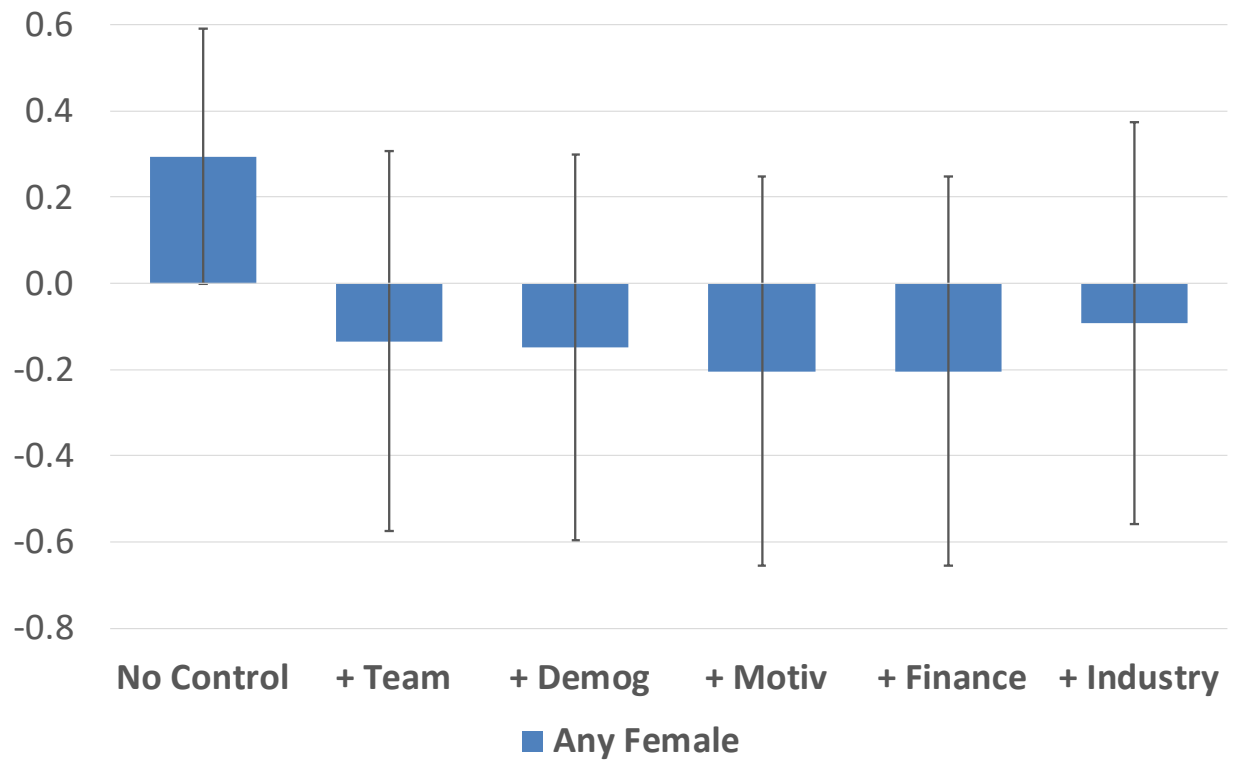**Figure 5a. Innovation Count (Any Female vs. All Male)**

Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430

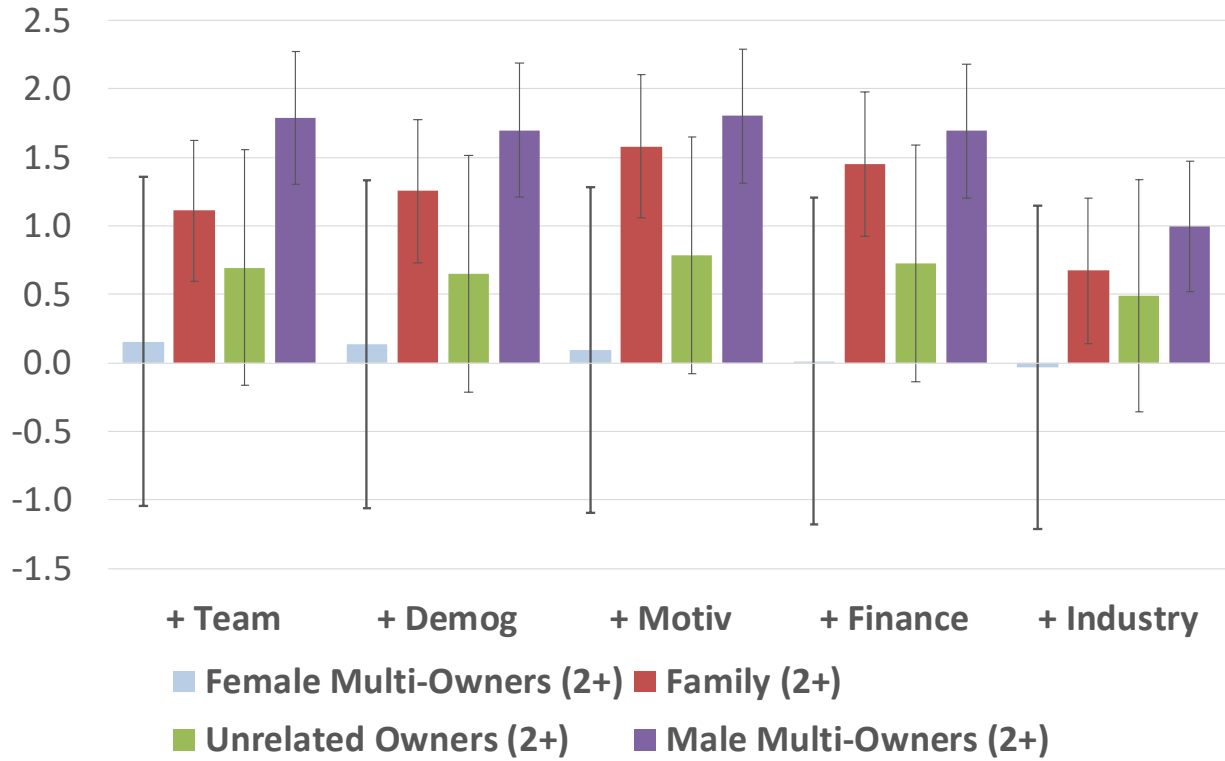**Figure 5b. Innovation Count (Disaggregated Categories vs. Single Male Owner)**

Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430

Figure 6a. Good/Service No Other Firm Offers (Any Female vs. All Male)
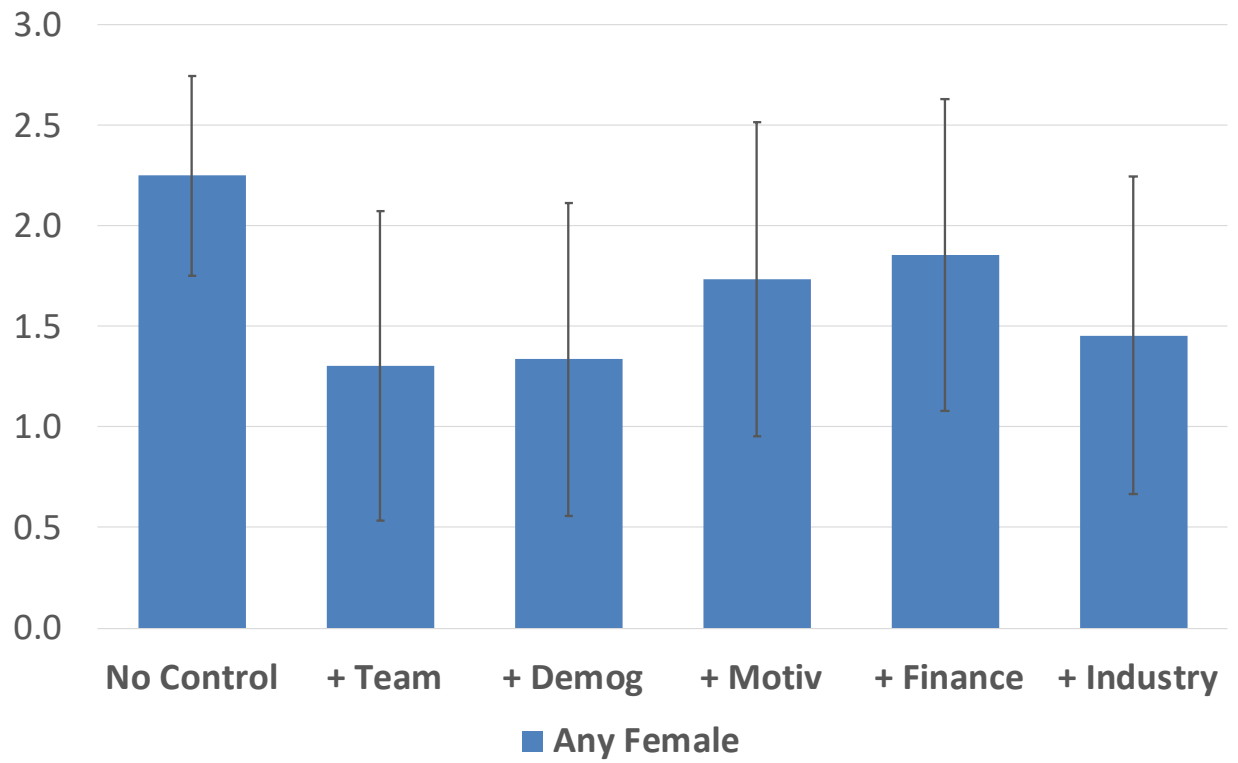

Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430

Figure 6b. Good/Service No Other Firm Offers (Disaggregated Categories vs. Single Male Owner)

Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430

Figure 7a. New Good/Service to this Firm (Any Female vs. All Male)

Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430

Figure 7b. New Good/Service to this Firm (Disaggregated Categories vs. Single Male Owner)

Figure 8a. Applied New Way of Support Activity (Any Female vs. All Male)

Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430

Figure 8b. Applied New Way of Support Activity (Disaggregated Categories vs. Single Male Owner)

Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430

Figure 9a. Decreased Production Costs (Any Female vs. All Male)

Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430

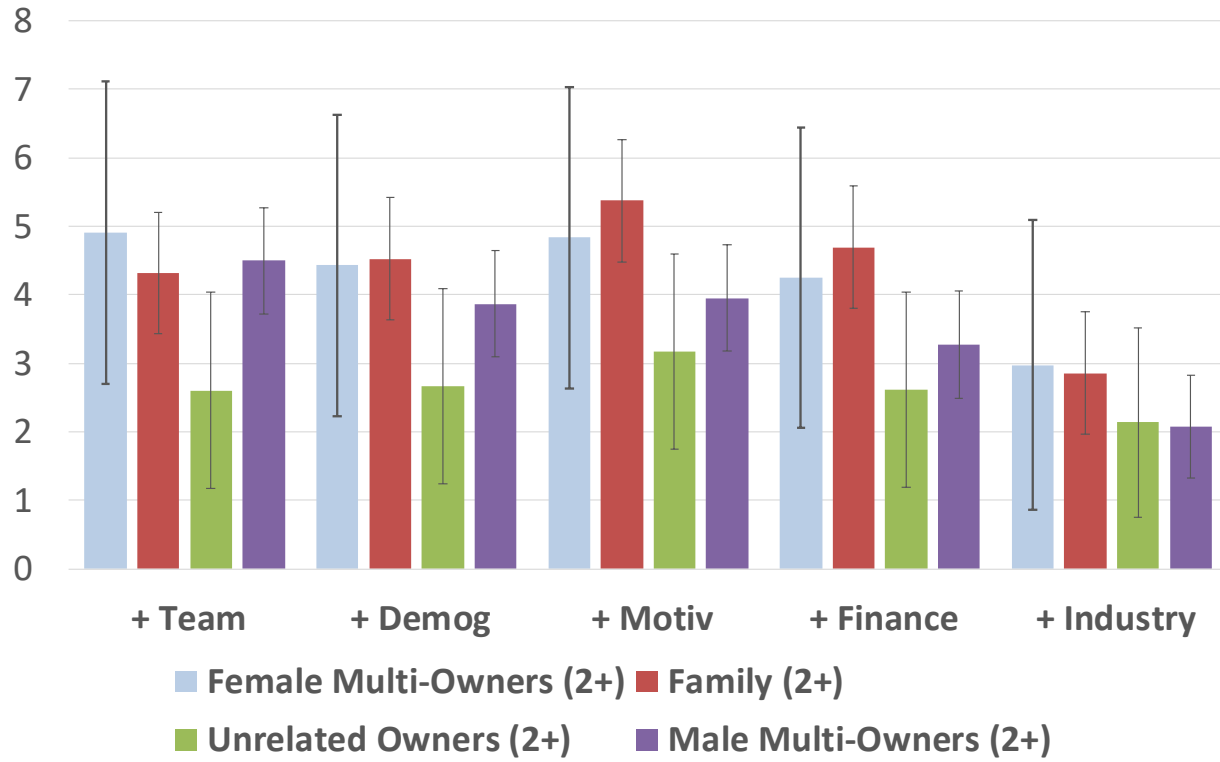**Figure 9b. Decreased Production Costs (Disaggregated Categories vs. Single Male Owner)**

Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430
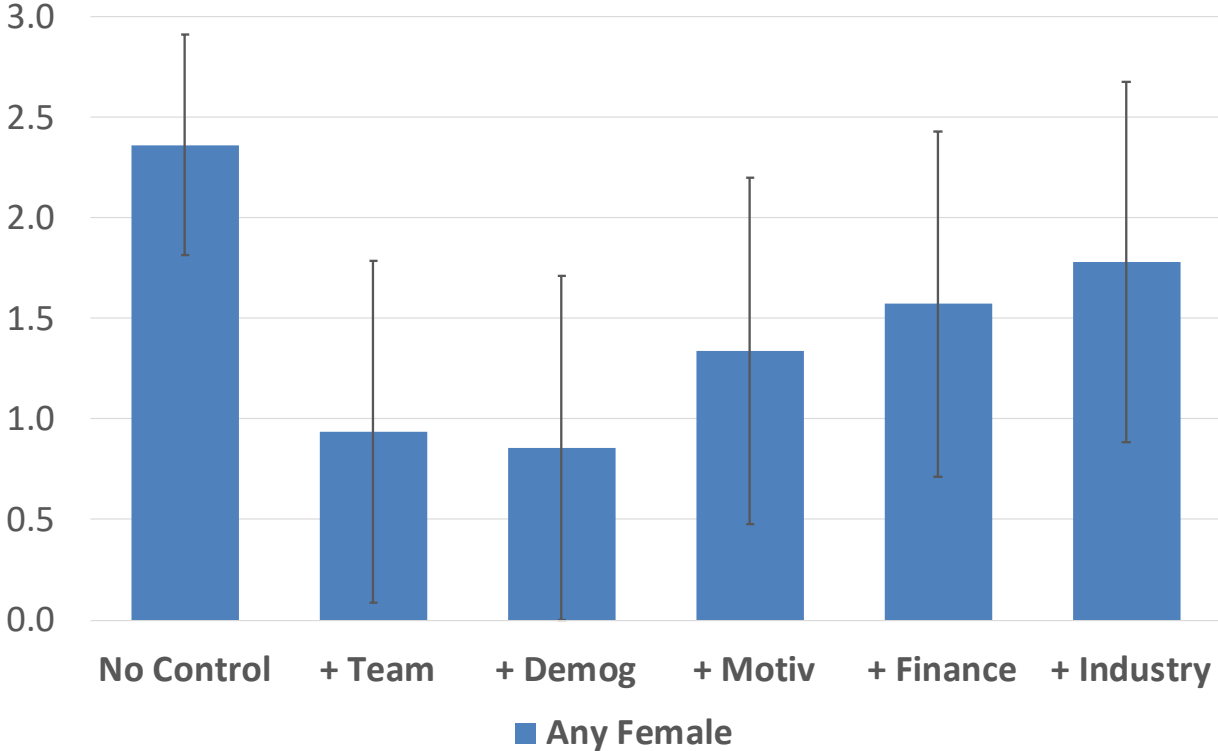
Figure 10a. Any R&D Activity (Any Female vs. All Male)

Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430

Figure 10b. Any R&D Activity (Disaggregated Categories vs. Single Male Owner)

Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430

Figure 11a. Number of Patents (Any Female vs. All Male)

Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430

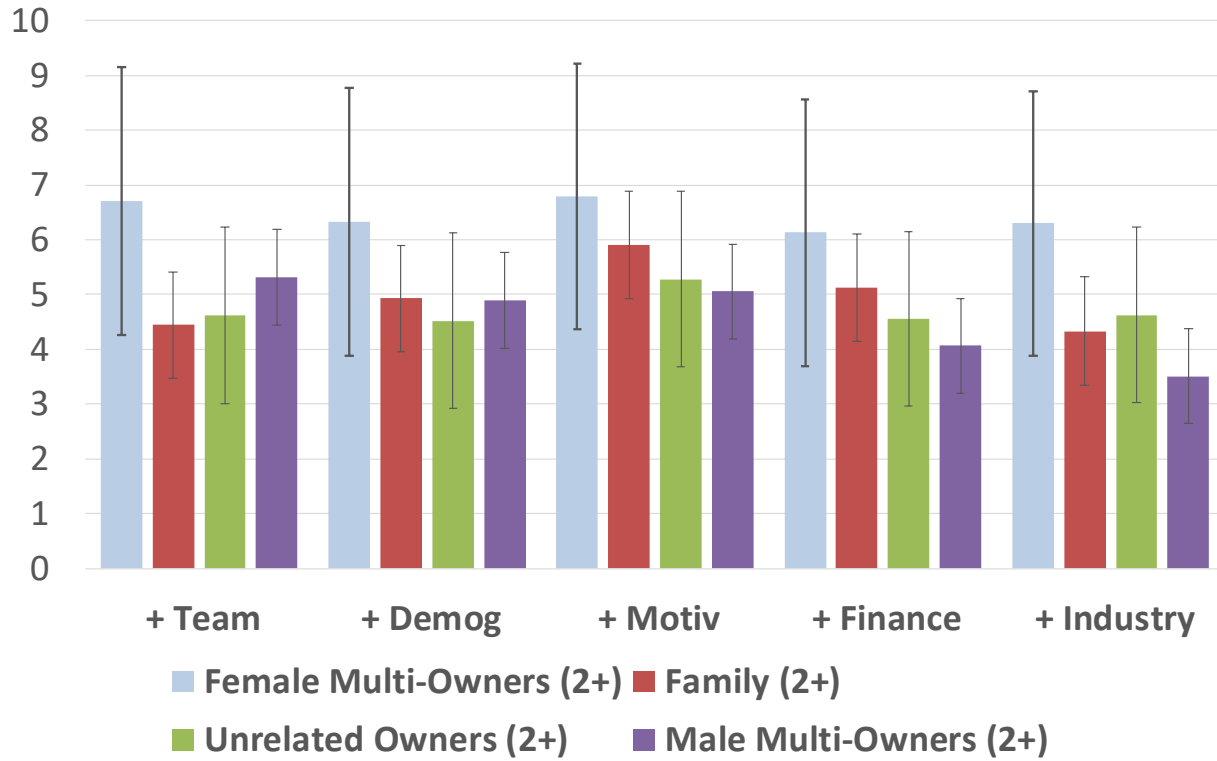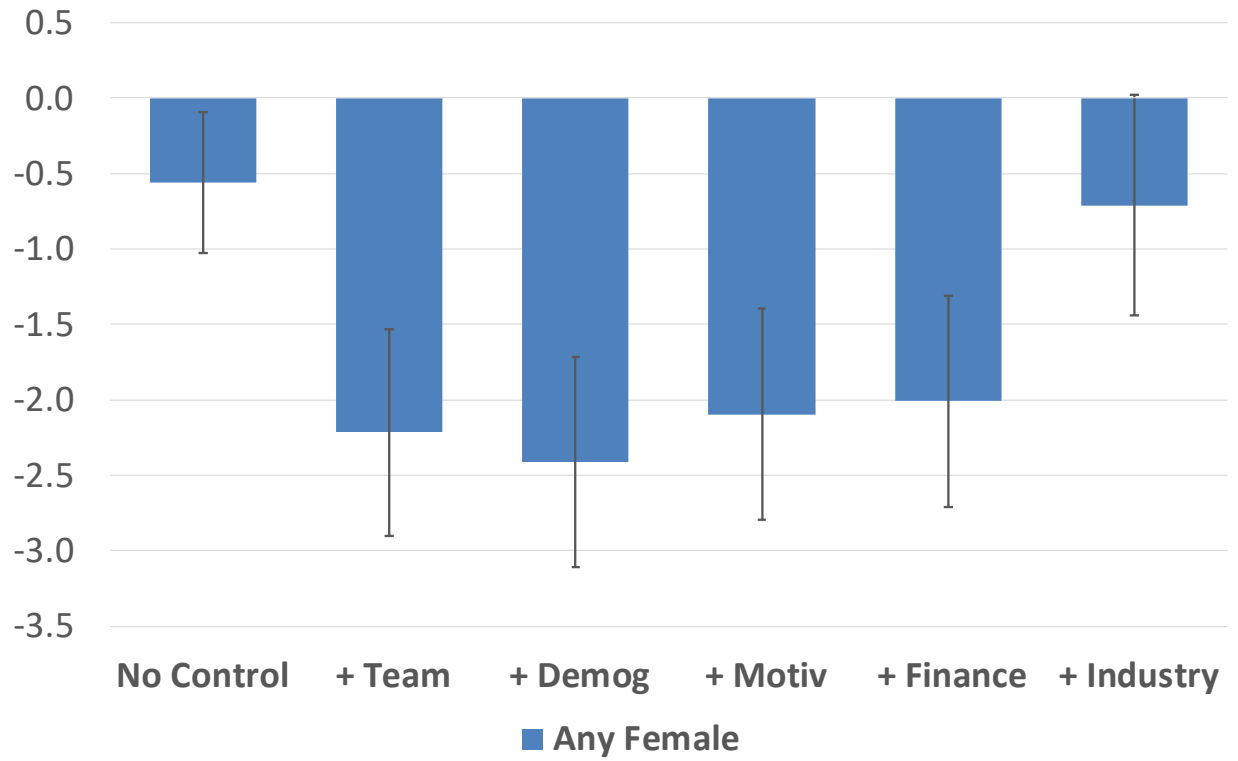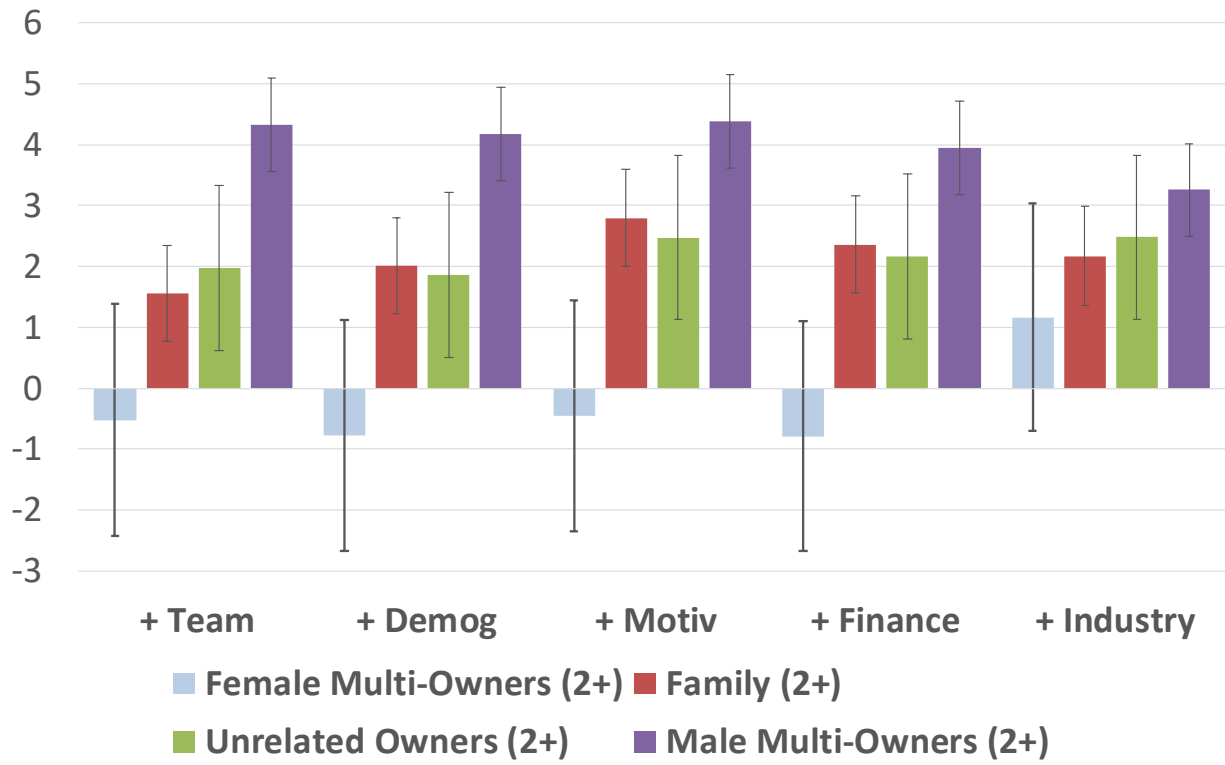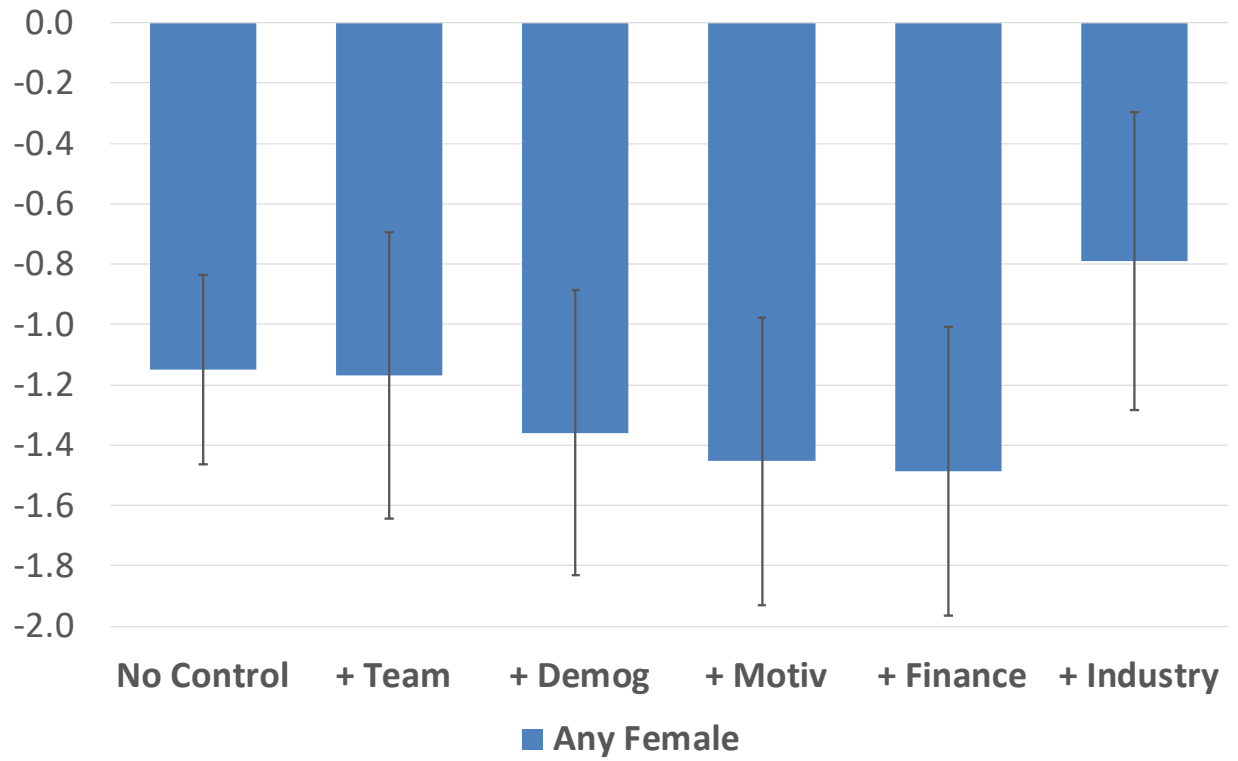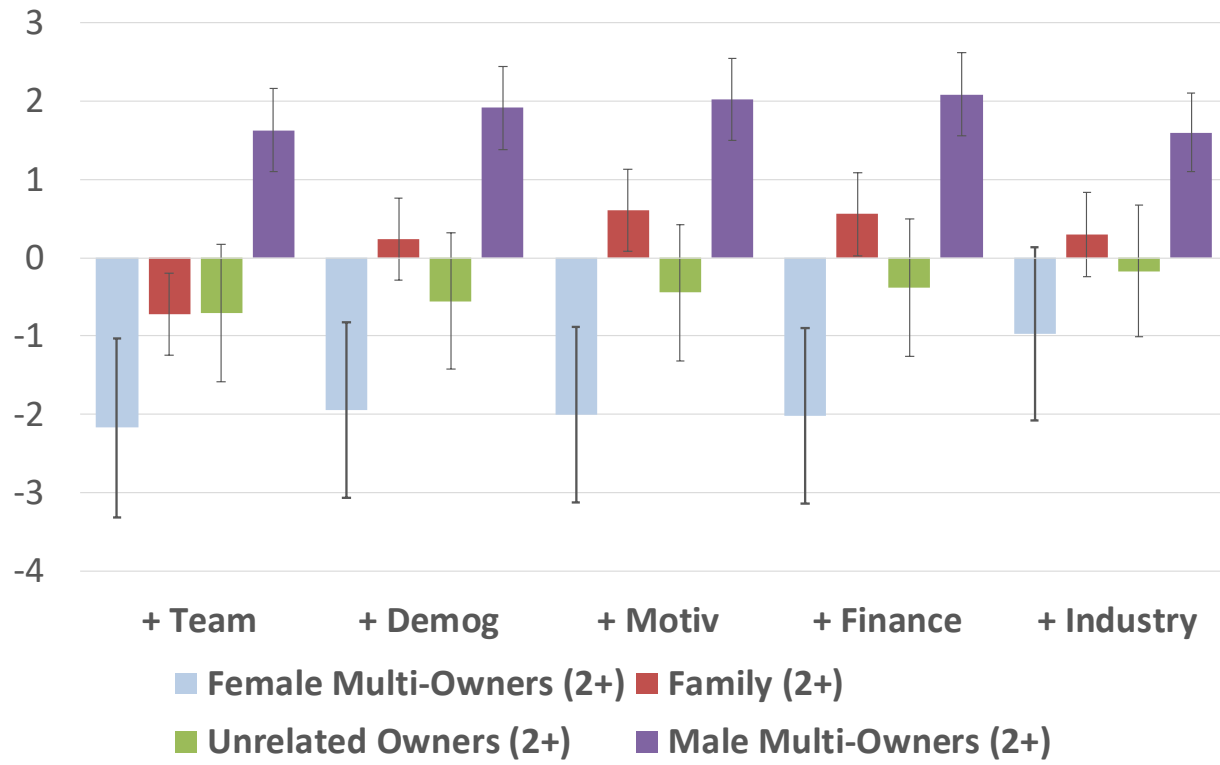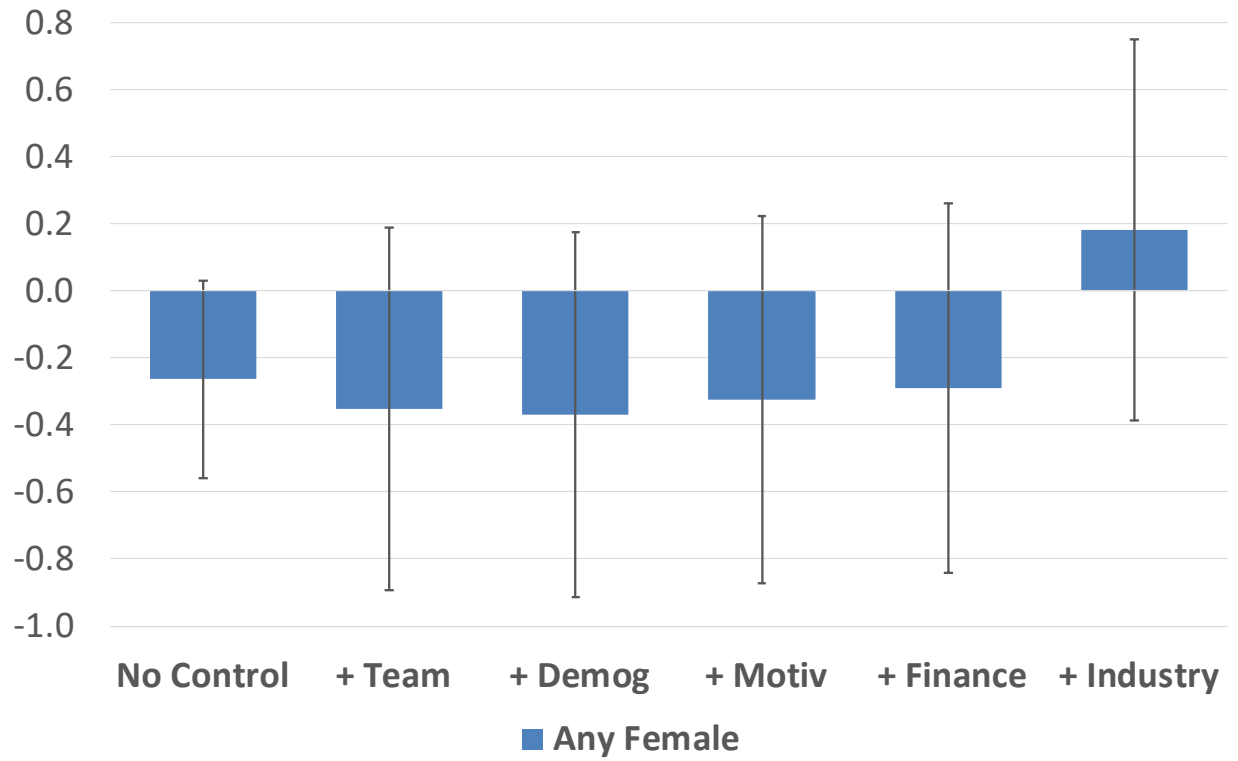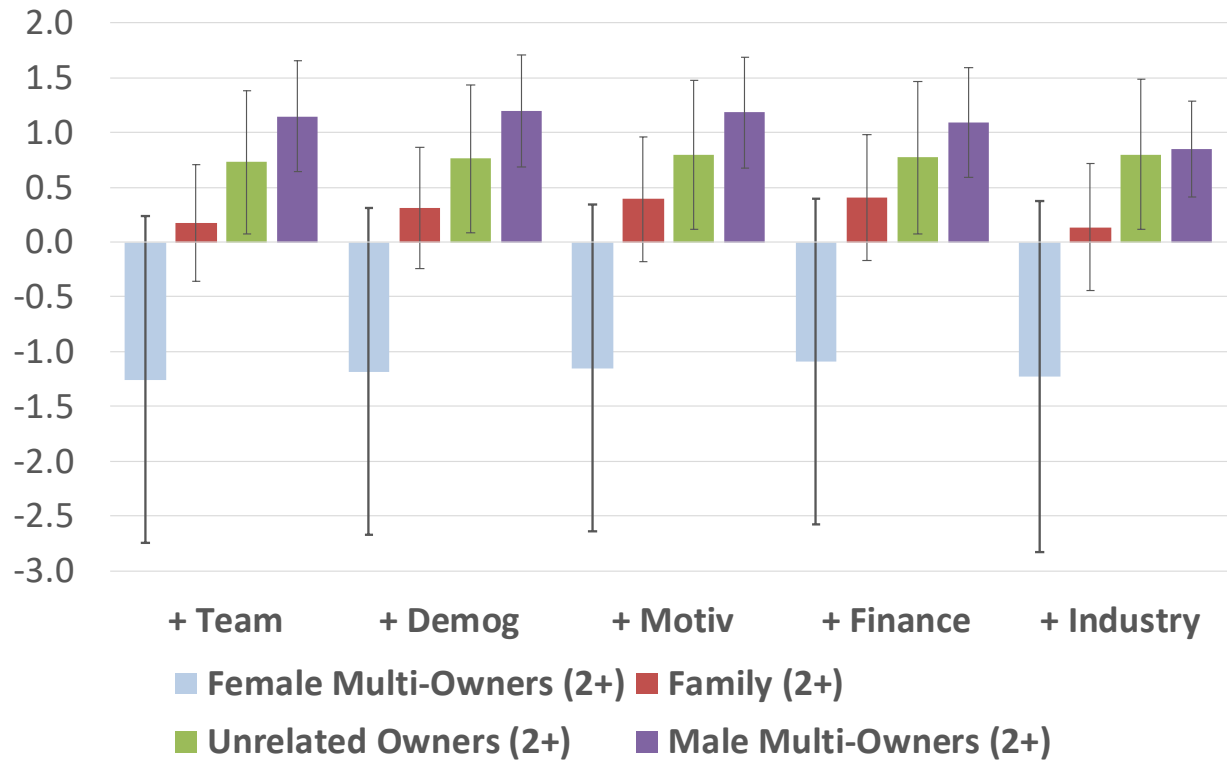**Figure 11b. Number of Patents (Disaggregated Categories vs. Single Male Owner)**

Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430

Table 1. Descriptive Statistics of Owner Characteristics

	All	Sub-group		
		All Female	**Diverse**	**No Female**
Ownership				
Single female	13.48	88.85	0.00	0.00
Single male	42.66	0.00	0.00	74.78
Family (2+)	23.24	0.00	83.67	0.00
Unrelated owners (2+)	4.54	0.00	16.33	0.00
Female multi-Owners (2+)	1.69	11.15	0.00	0.00
Male multi-Owners (2+)	14.39	0.00	0.00	25.22
Two or more owners	43.86	11.15	100.00	25.22
Demographics				
Age <25	0.19	0.17	0.17	0.20
Age 25-34	4.63	6.01	3.81	4.66
Age 35-44	16.53	18.80	15.39	16.48
Age 45-54	29.51	31.25	29.70	28.95
Age 55-65	31.52	29.51	32.59	31.53
Age 65 or over	17.63	14.27	18.33	18.18
White (non-Hispanic)	84.09	78.63	84.19	85.49
Black (non-Hispanic)	1.51	2.72	1.19	1.34
Asian (non-Hispanic)	8.40	11.78	8.30	7.55
Others (non-Hispanic)	0.79	0.87	0.88	0.74
Hispanic	5.21	6.01	5.43	4.89
Immigrant	15.53	18.02	14.90	15.18
Human Capital				
Less than high school	2.90	1.86	2.99	3.14
High school	18.06	14.55	20.54	17.78
Vocational school/Some college/Associate's degree	26.84	27.78	30.83	24.64
Bachelor's degree	27.62	26.11	28.55	27.56
Graduate degree	24.59	29.70	17.10	26.87
Veteran	10.16	1.21	8.87	13.16
Prior business	31.35	22.52	32.17	33.30
Observations (firm-owner)	179,000	17,000	67,000	95,000
Observations (firm)	119,000	15,000	33,000	71,000

Note: These are percentages of owners for all firms and for sub-groups (all female, diverse, and no female) by characteristics from the regression sample. Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430.

Table 2. Descriptive Statistics on Motivation

	All	Sub-group		
		All Female	**Diverse**	**No Female**
Motivation				
Own boss	60.89	58.32	52.83	65.50
Flexible hours	45.40	55.40	45.31	42.79
Balance work and family	50.68	57.74	51.95	48.18
Income	58.57	50.98	55.93	61.88
New idea	52.32	55.56	44.68	55.17
No job	2.79	3.44	2.41	2.80
Work for self	27.57	22.91	24.54	30.29
Always want to start business	44.10	38.82	38.19	48.39
Role model	24.39	25.04	24.12	24.36
Observations (firm-owner)	179,000	17,000	67,000	95,000
Observations (firm)	119,000	15,000	33,000	71,000

Note: The motivation variables are binary variable where 1 is very important and 0 otherwise. Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430.

Table 3. Descriptive Statistics on Finance

	All	Sub-group		
		All Female	Diverse	No Female
Finance				
No capital needed	7.76	9.58	6.61	7.84
Capital under $5k	14.01	18.25	12.15	13.79
5k to 10k	8.31	9.65	7.72	8.24
10k to 25k	12.27	12.91	11.90	12.29
25k to 50k	9.83	10.37	10.01	9.60
50k to 100k	10.88	10.18	12.00	10.52
100k to 250k	11.19	9.78	13.14	10.61
250k to 1m	7.56	5.64	9.11	7.32
1m to 3m	1.59	0.85	1.93	1.63
3m and more	0.49	0.19	0.55	0.55
Don't know start up capital	16.10	12.60	14.89	17.62
Startup funding from a bank or financial institution	20.35	16.57	23.08	20.02
Startup funding from venture capitalist(s)	0.51	0.23	0.51	0.58
Startup funding from gov't-guaranteed business loans	2.31	2.13	3.43	1.81
Startup funding from grants	0.18	0.20	0.19	0.16
Funding from banks or other financial institutions	37.00	31.93	40.57	36.61
Funding from investors	1.73	1.41	1.61	1.87
Applied additional funding from VCs and received total amount	0.11	0.08	0.07	0.13
Applied additional from banks and received total amount	10.51	7.61	11.20	10.95
Applied additional from angel investors and received total amount	0.15	0.13	0.10	0.18
Applied additional from other investor businesses and received total amount	0.22	0.25	0.11	0.26
Applied additional from crowdfunding and received total amount	0.08	0.11	0.07	0.07
Applied additional from grants and received total amount	0.19	0.27	0.23	0.15
Avoid additional funding Reason: Didn't think it would be approved by lender	4.10	5.02	4.62	3.60
Negative impact on profitability due to access to credit	9.64	9.67	10.60	9.17
Observations (firm-owner)	179,000	17,000	67,000	95,000
Observations (firm)	119,000	15,000	33,000	71,000

Note: These are percentages of owners by characteristics from the regression sample. Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430.

Table 4. Descriptive Statistics of Industry

	All	Sub-group		
		All Female	Diverse	No Female
Industry				
Primary sector	D	D	D	D
Construction	12.79	3.46	12.89	15.22
Manufacturing	5.15	2.47	6.49	5.20
Wholesale Trade	5.95	3.23	6.40	6.46
Retail Trade	11.90	11.77	14.60	10.61
Transportation	2.94	1.63	3.46	3.04
Information	D	D	D	D
Finance and Insurance	4.29	3.95	3.33	4.84
Real Estate	4.21	4.90	4.93	3.67
Business Services	16.08	18.89	12.59	17.04
Administrative Services	5.57	6.10	5.99	5.23
Educational Services	1.06	2.44	1.26	0.60
Health Care	12.35	21.42	7.33	12.38
Arts and Entertainment	1.38	1.66	1.45	1.27
Food Services	7.67	7.88	9.95	6.50
Other Services	7.00	9.11	7.42	6.23
Observations (firm-owner)	179,000	17,000	67,000	95,000
Observations (firm)	119,000	15,000	33,000	71,000

Note: "D" means suppressed to ensure that no confidential information is disclosed. Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430.

Table 5. Descriptive Statistics of Finance, Industry, and Other Choices

	All	Sub-group		
		All Female	**Diverse**	**No Female**
Other Choices				
Manager	81.11	86.11	70.91	84.74
Service Provider	63.75	71.51	51.87	67.47
Financial Controller	74.97	79.31	69.41	76.52
No role	4.54	2.77	8.56	3.05
No Hours Worked	3.80	2.28	6.74	2.77
Less than 20 hours Worked	11.73	9.46	18.25	9.16
20-39 hours Worked	15.03	19.46	16.90	12.93
40 hours Worked	15.74	17.54	14.62	15.81
41-59 hours Worked	32.80	31.73	26.68	36.07
60 hours or more Worked	20.89	19.52	16.81	23.24
Primary Income	76.89	77.37	69.25	80.48
Home-based	20.26	22.30	21.44	19.15
Observations (firm-owner)	179,000	17,000	67,000	95,000
Observations (firm)	119,000	15,000	33,000	71,000

Note: These are percentages of owners by characteristics from the regression sample. Source: ASE, Census DRB bypass number DRB-B0047-CED-20190430.

Table 6. Descriptive Statistics of Firm Performance

	All	Sub-group		
		All Female	Diverse	No Female
Innovation				
Innovation dummy (%)	57.35	56.36	60.46	56.11
Innovation count	2.35	2.22	2.52	2.31
Good/Service no other firm offers (%)	5.11	4.54	5.77	4.93
New good/service to this firm (%)	16.61	16.38	18.85	15.58
New way to support activity (%)	21.78	21.06	24.35	20.72
Decreased production costs (%)	14.52	11.61	15.66	14.73
Any R&D activity (%)	6.04	4.84	5.75	6.50
Number of patents	3.47	1.23	3.68	3.97
Employment				
Employment in 2014	12	8	13	13
Top 5 % employment in 2014 (38+ employees)	119	114	115	122
Bottom 95 % employment in 2014 (1-38 employees)	6	5	7	6
Aspiration to Grow				
Aspiration to be larger	67.31	64.37	68.53	67.5
Export				
Conducted exports outside the U.S.	5.96	3.75	6.95	6.07
Observations (firm-owner)	179,000	17,000	67,000	95,000
Observations (firm)	119,000	15,000	33,000	71,000

Note: These are percentages of owners by innovation measures (except innovation count and number of patents) from the regression sample. Aspiration to be larger equals to 1 if the owner would like the business to be larger in terms of sales or profits in five years. Export equals to 1 if the business's sales of goods and/or services are consisted of exports outside the U.S. in 2014. Census DRB bypass number DRB-B0047-CED-20190430.