Online Appendix:
The Disparate Impact of Up-or-Out Promotion Policy on Fertility Timing

Kyung H. Park — Wellesley College
Nayoung Rim — U.S. Naval Academy

May 4, 2019

Contents

1 Additional Descriptive Statistics

1.1 Comprehensive List of Firm Types in *After the JD* survey

- Private Law Firm
- Solo Practice
- Federal Government
- State or Local Government
- Educational Institution
- Legal Services or Public Defender
- Public Interest Organization
- Other Non-Profit
- Professional Service Firm
- Other Fortune 1000 Industry/Service
- Other Business/Industry
- Labor Union or Trade Association
- Military
- Legal Temporary Firm
- Insurance Company

1.2 Wave 1 Questions on Important Factors and Determinants of Respondent's Initial Career Decisions

38. Thinking about the principal types of settings in which lawyers work (e.g., government, large law firms, business), how important was each of the following factors in determining the sector in which you began your professional career? *(Exclude clerkships.) Check one box on each line.*

		Not At All Important						Extremely Important	NA
a.	Medium-to-long-term earning potential	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
b.	Substantive interest in a specific field of law	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
c.	Salary to pay off law school debts	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
d.	Availability of loan repayment assistance or loan forgiveness programs	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
e.	Opportunity to develop specific skills	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
f.	Potential to balance work and personal life	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
g.	Opportunity to do socially responsible work	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
h.	Prestige of the sector	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
i.	Opportunities for future career mobility	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
j.	Other (*Specify:* ______________________________)	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐

40. Comparing specific job offers you received from employers you considered, how important were the following factors in making your choice? *Check one box on each line.*

I received one offer ➔ **Skip to Question #41 on page 12.**

		Not At All Important						Extremely Important	NA
a.	Salary	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
b.	Benefits	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
c.	Office environment/collegiality	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
d.	Hours expected	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
e.	Pro bono opportunities	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
f.	Prospects for advancement	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
g.	Good match of employer's mission and my own	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
h.	Location	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
i.	Size	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
j.	Prestige	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐
k.	Training/mentorship opportunities	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	☐

65. How important was each of the following goals in your decision to attend law school?
Check one box on each line.

		Irrelevant				Very Important
a.	Intellectual challenge of law school and the law	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
b.	Desire to help individuals as a lawyer	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
c.	Desire to develop a satisfying career	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
d.	Desire to defer entry into the job market	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
e.	Desire for eventual financial security	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
f.	Desire to change or improve society	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
g.	Becoming influential in a powerful profession	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
h.	Desire to build a set of transferable skills	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
i.	Other (*Specify:* ________________________)	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

1.3 Sampling Process

In the first stage of the sampling process, the nation was divided into 18 strata by region and size of the new lawyer population. Within each stratum, one primary sampling unit (PSU) was selected. A PSU can be a metropolitan area, portion of a state outside large metropolitan areas, or the entire state. The PSUs included the four "major" markets, those with more than 2,000 new lawyers (Chicago, Los Angeles, New York, and Washington, D.C.); 5 of the 9 "large" markets, those with between 750-2,000 new lawyers; and nine of the remaining "smaller" markets. In the second stage, researchers sampled individuals from each of the PSUs from databases of individuals addmitted to a bar in 2000. This made up 7,727 individuals. An oversample of 1,465 minority lawyers (Black, Hispanic, and Asian American) yielded a final sample of 9,192 lawyers.

However, 20 percent of the sample could not be located and another 20 percent of those located were identified as lawyers moving from one state bar to another rather than having entered the bar for the first time in 2000. These "movers" were left in the sample as long as they graduated from law school no earlier than 1998. The response rate for Wave 1 was 71 percent, yielding an analysis sample of 4,538 lawyers.

1.4 Sample Attrition

In this section, we discuss attrition from the AJD survey. First, we focus on the possibility that our results reflect endogenous attrition rather than a response to up-or-out policy. For example, if women with children disappear at faster rates, then women who remain in the AJD will have higher rates of childlessness than men due to attrition. Second, we explain our decision to apply a re-weighting procedure in greater detail. Although attrition does not systematically relate to fertility timing, our final sample of stayers differs modestly from the initial sample in Wave I along some margins. This raises a potential concern that our results are not representative of the original population of interest and instead specific to respondents who stay in the AJD.

In summary, we do not find that women with children exit the AJD at higher rates. We also show that our results are highly robust to dropping weights altogether and to various approaches to estimating the weights. This analysis mitigates concerns that attrition leads to a sample of attorneys that are not broadly representative of the legal profession.

Endogenous Attrition

A primary concern is that selective attrition could lead to biased estimates of the gender difference in fertility timing. Suppose, for example, that women who participate in all three waves of the AJD have a higher propensity to have children in late- versus early-career and that women who are more likely to exit the AJD are also those who tend to have children in early-career. In this case, women could, in fact, have similar fertility timing as men but appear to delay because women who have children in early-career are less likely to report to the AJD. This does not seem implausible especially if women invest more time in child care than men and child care crowds out time spent on extraneous activities including participation in the AJD. We examine whether or not women with children are more likely to exit the AJD prior to completion next.

Table A1 shows descriptive statistics for three different samples; (i) respondents who are in Wave I but not Wave III, (ii) respondents who are in Waves I and II but not Wave III, and (iii) respondents who are both in Waves I and III. We refer to the first two subgroups as *Wave I leavers* and *Wave II leavers*, respectively, and the third as *stayers*. Panel A focuses on gender composition and fertility outcomes. There is no evidence that women exit the AJD at higher rates than men; in fact, women actually constitute a minority among those who leave the AJD early (42% and 39.8% among Waves I and II leavers, respectively) and roughly half of those who stay (50.4%). There is also no evidence that women with children are more likely to exit as the fraction of Wave I and Wave II leavers that are both female and have children is 0.094 and 0.085, respectively, which is less than the 0.100 value among stayers. This implies that women with children are slightly more likely to stay rather than leave the AJD which is inconsistent with the narrative that women with children disappear at high rates. Finally, age and marital status are also fairly even across all three groups which is noteworthy because these demographic characteristics are important determinants of fertility timing. Overall, these statistics are inconsistent with the idea that our results are driven by endogenous attrition in the AJD.

In Panel B, we present additional descriptive means along other margins. There are some substantive differences worth noting. For example, roughly 7.3% of stayers are African-American whereas the analogous percentage among Wave I leavers is 9.1%, and this 1.8 percentage point difference is statistically significant at the 5% level. Stayers are also slightly less likely than Wave I leavers to be in a private law firm. However, many of these variables show balance across the three subgroups. To take one example, the average salary for Wave I and Wave II leavers is $86,615 and $85,430, respectively, which is similar to the $83,961 average salary among stayers. Further, there are only minor differences in debt, hours worked, firm size, fraction Hispanic, Asian, and residing in NY, LA, Chicago or DC across the three groups. Taken together, Panels A and B suggest that attrition is unlikely to generate bias as the differences between stayers and leavers are largely unrelated to fertility and are fairly modest along other non-fertility related margins.

Table A1: Descriptive Statistics of Leavers vs. Stayers

Panel A: Gender and Fertility Outcomes	Wave I Leavers			Wave II Leavers			Stayers		
	Mean	Std Dev	N	Mean	Std Dev	N	Mean	Std Dev	N
Female	0.421	0.494	1,895	0.398	0.49	832	0.504	0.5	2,460
Birth year	1970.226	5.859	1,865	1,970.52	5.801	820	1970.459	5.561	2,434
Ever married	0.603	0.489	1,878	0.601	0.49	825	0.623	0.485	2,436
Parent	0.273	0.445	1,860	0.248	0.432	819	0.248	0.432	2,408
Female Parent	0.094	0.291	1,857	0.085	0.28	819	0.1	0.3	2,401

Panel B: Additional Characteristics	Wave I Leavers			Wave II Leavers			Stayers		
	Mean	Std Dev	N	Mean	Std Dev	N	Mean	Std Dev	N
Black	0.091	0.288	1,859	0.077	0.267	819	0.073	0.26	2,421
Hispanic	0.085	0.279	1,859	0.082	0.274	819	0.081	0.272	2,421
Asian	0.09	0.286	1,859	0.082	0.274	819	0.084	0.277	2,421
Major region	0.375	0.484	2,019	0.366	0.482	858	0.351	0.477	2,518
Undergraduate GPA	-0.037	0.036	818	-0.059	0.050	415	0.010	0.025	1,683
Law school GPA	-0.050	0.042	596	-0.036	0.049	375	0.005	0.025	1,657
Law school Tier	3.322	1.252	1,879	3.25	1.232	828	3.133	1.226	2,448
Private Law	0.656	0.475	1,877	0.665	0.472	814	0.628	0.484	2,399
Firm size	221.378	417.036	1,567	203.577	371.935	691	210.926	369.4	2,080
Salary	86,615.04	50,047.64	1,649	85,430.67	43,613.90	729	83,961.11	47,248.72	2,244
Hours worked	48.141	14.908	1,630	48.464	14.171	696	48.01	13.542	2,013
Debt	61,635.57	42,180.02	1,784	61,271.86	42,011.90	789	62,115.99	41,109.68	2,377

Source: AJD restricted data.
Notes: Major region includes NY, LA, Chicago, and DC. Undergraduate and law school GPAs are standardized to have mean 0 and standard deviation 1.

Inverse Probability Weighting

Nonetheless, there is a second-order concern associated with attrition. In particular, our final sample may no longer be representative of our original population of interest due to the modest differences between stayers and the initial Wave I sample (e.g. stayers are slightly more likely to be female and less likely to be African-American than leavers). A common technique to correct for this is to adjust the sampling weights using the familiar inverse probability weighting approach. This requires the researcher to (i) estimate the probability of staying conditional on observable characteristics and then (ii) divide the initial sampling weights by the estimated conditional probability of staying in the AJD.[1] This effectively re-weights the data such that respondents who stay but resemble those who leave the AJD

[1]Formally, we can think of problem as one in which the empirical object of interest is $E[Y|X]$ but we observe only $E[SY|X]$ where S is indicator variable for whether or not the respondent stays in the AJD. We can rewrite the latter term as $E[SY|X] = E[Y|S = 1, X]P(S = 1|X)$. If the outcome of interest is statistically independent from S conditional on X such that $E[Y|S = 1, X] = E[Y|S = 0, X]$, then we can drop S from the conditioning set (i.e. $E[Y|S = 1, X] = E[Y|X]$). The descriptive means in Panel A are strongly consistent with this *conditional independence* assumption. Then, re-weighting the data by $\frac{1}{P(S=1|X)}$ will allow the researcher to recover $E[Y|X]$.

prior to Wave III along observable characteristics receive greater weight. The re-weighted covariate distribution will then match the original once the weights are applied. Thus, our motivation for the re-weighting procedure is to ensure that our results represent the original population of interest.

Table A2 shows that our results are insensitive to different approaches to estimating the weights. To facilitate comparison, we present the results reported in the paper in column (1). The estimates imply that female lawyers are about 7.9 percentage-points less likely and 5.8 percentage-points more likely to have first birth in early-career and late-career, respectively, than males. Column (2) shows estimates without re-weighting the data. Columns (3) to (5) shows results in which we perturb the specification used to estimate the conditional probability of staying, the key input into the weights. The details of the specification are denoted in the bottom panel. Importantly, the estimates are fairly similar across these five specifications. For example, the gender difference in early parenthood lies within a tight range between -7.6 and -9.5 percentage-points and the respective range for late parenthood is 5.5 and 7.2 percentage-points. In addition, all of these estimates are statistically significant at either the 5% or 1% level. Thus, our results are fairly stable with and without weights and across different specifications used to construct the weights.

Table A2: Gender Difference in Early Parenthood and Late Parenthood: Attrition Checks

Specification:	Original (1)	No weights (2)	Original + Attr. Wts 1 (3)	Original + Attr. Wts 2 (4)	Original + Attr. Wts 3 (5)
Panel A: Early parenthood					
Female	-0.0789***	-0.0778***	-0.0776***	-0.0832***	-0.0969***
	(0.0245)	(0.0203)	(0.0245)	(0.0241)	(0.0234)
Observations	2,087	2,087	2,087	2,087	2,087
Adjusted R-squared	0.294	0.282	0.293	0.298	0.297
Panel B: Late parenthood					
Female	0.0581**	0.0663***	0.0567**	0.0608**	0.0745***
	(0.0247)	(0.0204)	(0.0247)	(0.0243)	(0.0237)
Observations	2,087	2,087	2,087	2,087	2,087
Adjusted R-squared	0.302	0.286	0.300	0.308	0.303
Main Regression Controls:					
Demographic characteristics	Yes	Yes	Yes	Yes	Yes
Ability proxies	Yes	Yes	Yes	Yes	Yes
Job characteristics	Yes	Yes	Yes	Yes	Yes
Income and spousal employment	Yes	Yes	Yes	Yes	Yes
Factor scores	Yes	Yes	Yes	Yes	Yes
Attrition Model Specification:					
Functional Form	Probit	None	LPM	Probit	Probit
Interactions	Female		Female	None	Female
Work-related controls only					Yes

Source: AJD restricted data.
Notes: Likelihood of early parenthood is 0.51 (males) and 0.40 (females). Likelihood of late parenthood is 0.44 (males) and 0.53 (females). Early parenthood is defined as having one's first child within 7 years after the JD. Late parenthood is defined as having one's first child at least 9 years after JD. Childless lawyers are classified as having a late parenthood. Demographic characteristics account for race and ethnicity, age, law school graduation date, geographic location at time of initial survey, and initial marital status. Ability proxies include standardized undergraduate and law school GPAs, U.S. News' 2003 law school ranking, participation in general law review, judicial clerkships, number of initial job offers, and number of bar exam attempts. Job characteristics control for initial firm's size, early-career weekly hours worked, and area of law. Employment controls include spousal employment status at time of initial survey, early-career household income, and respondent's early-career salary. Factor scores are: social responsibility, earning potential, prestige, career development, firm's ranking, mission match, and financial security. *** p< 0.01, ** p< 0.05, * p< 0.1

1.5 Imputation of Spousal Income Trajectory

We predict the spousal income trajectory using the three waves of cross-sectional data on spouse's income. However, for respondents who reported spousal income in only one wave or did not report income for an employed spouse, we use Census and ACS data to impute the spouse's income trajectory (288 respondents). This section describes our imputation methodology.

We use the ACS data from years 2001-2014 and 1990 and 2000 Census data. We keep only adults aged between 17 and 64, inclusive. We drop those living in group quarters and unmarried couples. That is, we keep married couples where at least spouse is in the legal profession (OCC2010 code 2100). We also double-check the lawyer's education by dropping those who do not have a graduate degree.

After adjusting the top-coded income responses (we multiply the top-coded value by 1.5), we calculate hourly wages and trim extreme wages (dropping wages that are between \$0-\$1 and greater than 1/35 of the annual maximum weekly wage.

We predict the spousal income using the lawyer's gender, race and ethnicity, and years of experience. Years of experience is defined as: $\max\{\texttt{Age} - 19 - 5, 0\}$. The reason we have a sparse specification is that the AJD survey has very little information on the spouse; we only know the spouse's employment status and income. We run the following OLS regression model on logged spousal hourly wage:

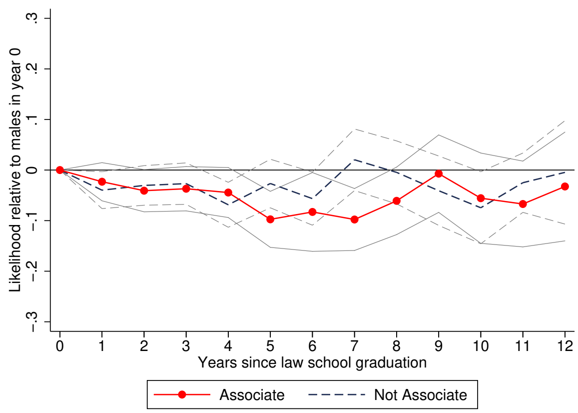
$$\begin{aligned} \texttt{log(spousal hrly wage)} &= \beta_0 + \beta_1 \cdot F_i + \beta_2 \cdot R_i + \beta_3 \cdot \Big(F_i \times R_i\Big) \\ &+ \delta_1 \cdot exp_i + \delta_2 \cdot exp_i^2 + \delta_3 \cdot exp_i^3 + \delta_4 \cdot exp_i^4 \\ &+ \delta_5 \cdot \Big(F_i \times exp_i\Big) + \delta_6 \cdot \Big(F_i \times exp_i^2\Big) \\ &+ \delta_7 \cdot \Big(F_i \times exp_i^3\Big) + \delta_8 \cdot \Big(F_i \times exp_i^4\Big) \\ +\varepsilon_i \end{aligned} \tag{1}$$

These coefficients are then used to predict spousal income in the AJD data. For respondents who reported spousal income in Wave 1 (2001), we back out what it would have been in the year that the respondent graduated from law school, which was not necessarily in 2001.

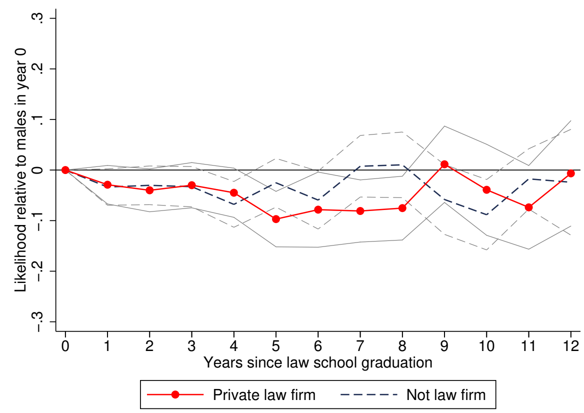
2 Additional Results

Figure A1: Gender Difference in Hazard Rate by Alternative Measures of Exposure to Up-or-Out Policies

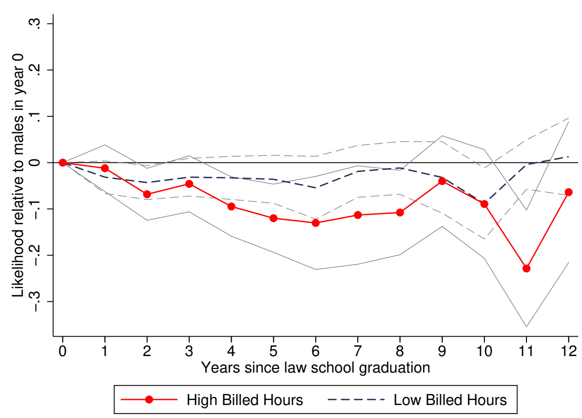
(a) Associate



(b) Private Law Firm

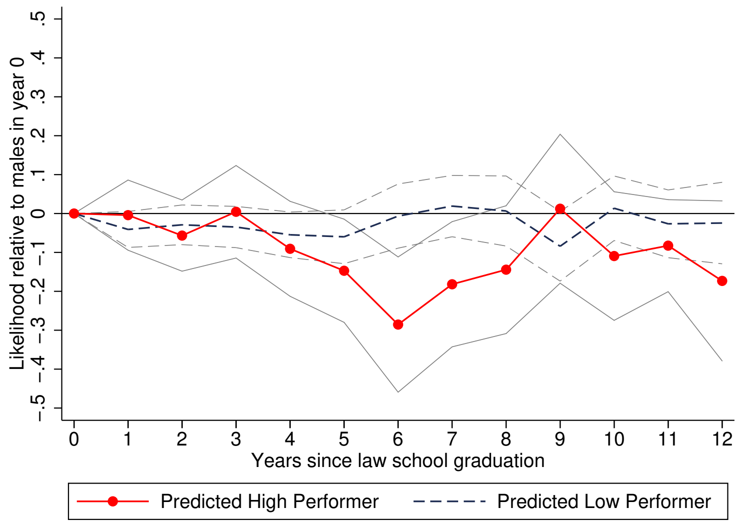


(c) Billed Hours in Mid-Career



Source: AJD restricted data.
Notes: N = 18,081. This figure depicts each year's contribution to the female's hazard of exiting the childless state relative to male's in year 0, separately by exposure measure. Associates and Private Law Firms refer to the lawyer's first job after law school. High billed hours are defined as 1,561 or more annual billed hours. Gray lines are 95% confidence intervals.

Figure A2: Gender Difference in Hazard Rate Among Married Lawyers

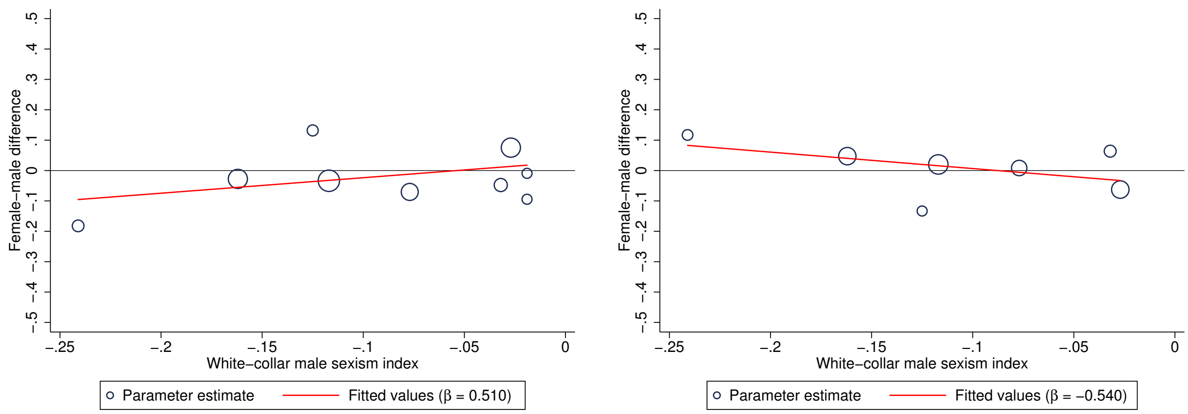


Source: AJD restricted data.
Notes: N = 9,297. This figure depicts each year's contribution to the female's hazard of exiting the childless state relative to male's in year 0, separately for high-intensity and low-intensity lawyers. Sample is restricted to those who reported being ever-married in Wave 1. Lawyers are classified as high performers if their predicted performance-level is in the top quartile and as low performers otherwise. Gray lines are 95% confidence intervals.

Figure A3: Gender Fertility Difference Among Low-Performing Lawyers by Level of Gender Norms

(a) Early Parenthood

(b) Late Parenthood



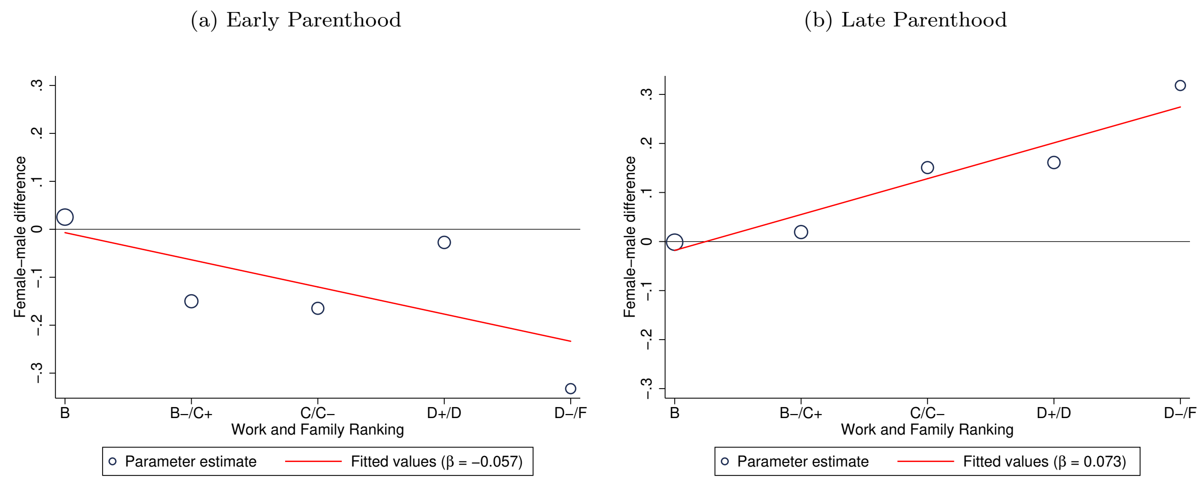
Source: AJD restricted data, Pan (2015).
Notes: These figures plot region-specific gender differences in early parenthood and late parenthood by the level of gender norms. The level of gender norms is classified by the Census Region's white-collar male sexism index, which is constructed from the GSS survey (Pan, 2015). More positive values correspond with more gender-prejudiced attitudes. Size of circles represents the census region's population. Fitted values are weighted by population size.

Table A3: Census regions by white-collar male sexism index

Census region	Sexism index
E. South Central (Tennessee)	-0.019
S. Atlantic (DC, Atlanta, Florida)	-0.027
W. South Central (Houston, Oklahoma)	-0.032
Mountain (Utah)	-0.06
Middle Atlantic (New York City, New Jersey)	-0.083
Pacific (Los Angeles, San Francisco, Oregon)	-0.117
W. North Central (Minneapolis)	-0.125
E. North Central (Chicago, Indiana, St. Louis)	-0.162
New England (Boston, Connecticut)	-0.241

Source: Pan (2015), Table 5.
Notes: Sexism index is constructed using white-collar male responses from the 1977-1998 GSS data. Positive values depict more sexist attitudes.

Figure A4: Gender Fertility Difference by Family-Friendly Work Laws



Source: AJD restricted data, Institute for Women's Policy Research's Status of Women 2015 report.
Notes: These figures plot region-specific gender differences in early parenthood and late parenthood by family-friendly work laws. Work and family conditions are classified by the state's grade in the Institute for Women's Policy Research's Status of Women 2015 report. Grades are assigned using paid leave legislation, elder and dependent care, child care, and the gender gap in parents' labor force participation rates. The regions in each grade category are: New York City, DC, Los Angeles, and San Francisco (B); New Jersey, Oregon, Minneapolis, Oklahoma (B-/C+); Atlanta, Connecticut, St. Louis, Boston (C/C-); Chicago, Houston, Tennessee (D+/D); Florida, Indiana, Utah (D-/F). Size of circles represents the region's sample size. Fitted values are weighted by region size.

Table A4: Fertility Outcomes Among High-Skilled Occupations

	Age at first child		Share parent		N	
Occupation	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**
Economists and market researchers	30.7	29.5	60%	55%	1,596	1,540
Chief executives	30.3	29.8	76%	58%	32,225	9,367
Lawyers and judges	31.9	32.1	77%	63%	20,515	10,096
Post-secondary Teachers	32.3	30.2	64%	63%	13,428	15,870
Accountants and Auditors	30.4	28.4	68%	67%	20,965	32,970
Pharmacists	30.7	29.6	77%	72%	3,206	2,862
Physicians and Surgeons	31.8	32.2	80%	72%	17,686	6,320
Everyone	29.3	27.7	59%	67%	2,200,000	2,200,000

Source: 2000 Census.

Notes: The sample is restricted to 44 year olds. Respondents are classified as parents if there is at least one child living in their household. N is the number of individuals. Everyone includes the universe of occupations.

Table A5: Gender Selection into Initial Job

	Outcome Variable: First Job at Private Law Firm						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.0478	-0.0412	-0.0412	0.00379	-0.0180	-0.0287	0.00723
	(0.0299)	(0.0286)	(0.0286)	(0.0277)	(0.0285)	(0.0278)	(0.0276)
Observations	1,780	1,780	1,780	1,780	1,780	1,780	1,780
R-squared	0.002	0.098	0.098	0.215	0.146	0.159	0.270
Controls for:							
Demographic characteristics		Yes	Yes	Yes	Yes	Yes	Yes
Ability proxies			Yes	Yes	Yes	Yes	Yes
Why Sector				Yes			Yes
Why Law					Yes		Yes
Why Job						Yes	Yes

Source: AJD restricted data.

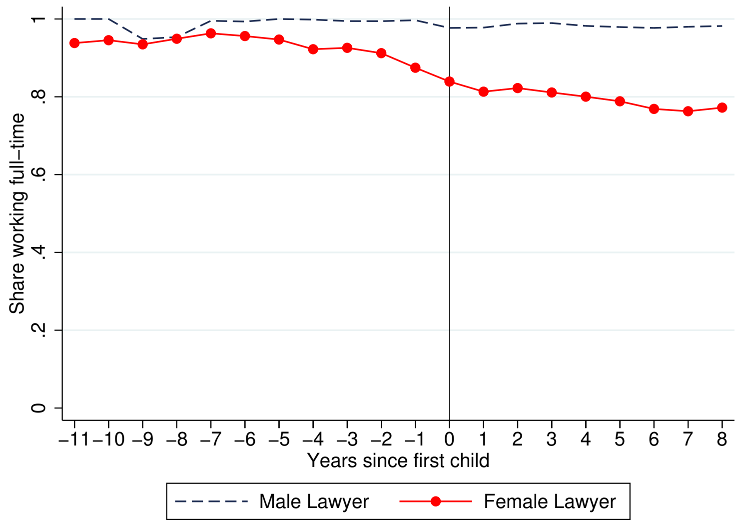
Notes: Demographic characteristics account for race and ethnicity, age, law school graduation date, geographic location at time of initial survey, and initial marital status. Ability proxies include standardized undergraduate and law school GPAs, U.S. News' 2003 law school ranking, participation in general law review, judicial clerkships, and number of bar exam attempts. Spousal employment and income controls include spousal employment status at time of initial survey, and early-career household income. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A6: Gender Fertility Difference by Predicted Performance

Predicted Performance:	High (1)	Low (2)
Panel A: Early parenthood		
Female	-0.147***	-0.0515*
	(0.0569)	(0.0274)
Panel B: Late parenthood		
Female	0.119**	0.0387
	(0.0577)	(0.0273)
Observations	2,087	2,087
Baseline controls	Yes	Yes

Source: AJD restricted data.
Notes: Early parenthood is defined as having one's first child within the first 7 years after law school. Late parenthood is defined as having one's first child at least 9 years after law school. Lawyers are classified as high performers if their predicted performance-level is in the top quartile and as low performers otherwise. See notes in Table A2 for description of baseline controls. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Figure A5: Full-Time Status by Birth of First Child



Source: AJD restricted data.
Notes: This figure depicts full-time status of male and female lawyers by years since their first child.

Table A7: Gender Difference in Promotion Probability

	Outcome Variable: Equity Partner						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.101***	-0.0906***	-0.0838***	-0.0729***	-0.0701***	-0.0462**	-0.0366*
	(0.0232)	(0.0227)	(0.0212)	(0.0209)	(0.0209)	(0.0201)	(0.0200)
Observations	1,780	1,780	1,780	1,780	1,780	1,780	1,780
R-squared	0.018	0.085	0.121	0.176	0.178	0.241	0.243
Controls for:							
Demographics		Yes	Yes	Yes	Yes	Yes	Yes
Ability proxies			Yes	Yes	Yes	Yes	Yes
Job characteristics				Yes	Yes	Yes	Yes
Income and spousal employment					Yes	Yes	Yes
Billed hours and caseload						Yes	Yes
Parental leave							Yes

Source: AJD restricted data.

Notes: See notes in Table A2 for description of demographics, ability proxies, job characteristics, and income and spousal employment controls. Parental leave controls include whether or not the lawyer took parental leave for his or her first child and the number of weeks taken. *** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

Table A8: Gender difference in fertility timing: Robustness Checks

Specification:	Alternative Thresholds (1)	Gender-Specific Factors (2)	Yulized Factors (3)	Survey Questions (4)
	Panel A: Early parenthood			
Female	-0.0675***	-0.0849***	-0.0849***	-0.0924***
	(0.0232)	(0.0247)	(0.0247)	(0.0247)
Observations	2,087	2,087	2,087	2,087
Adjusted R-squared	0.270	0.334	0.334	0.365
	Panel B: Late parenthood			
Female	0.0864***	0.0622**	0.0622**	0.0697***
	(0.0238)	(0.0250)	(0.0250)	(0.0249)
Observations	2,087	2,087	2,087	2,087
Adjusted R-squared	0.306	0.340	0.340	0.371
Main Regression Controls:				
Demographic characteristics	Yes	Yes	Yes	Yes
Ability proxies	Yes	Yes	Yes	Yes
Job characteristics	Yes	Yes	Yes	Yes
Income and spousal employment	Yes	Yes	Yes	Yes
Factor scores	Yes	Yes	Yes	
Why sector				Yes
Why law				Yes
Why job				Yes

Notes: Early parenthood is defined as having one's first child within the first 7 years after law school and late parenthood as having one's first child at least 7 years after law school. The alternative threshold specification defines "early parenthood" as having one's first child within the first 5 years after law school and "late parenthood" as having one's first child at least 7 years after law school (see Table A10). Gender-specific factors are ambition and family. Yulized factors are constructed from residuals of the regression of survey questions against female, race/ethnicity, marital status, number of children, undergraduate and law school GPAs, participation in general law review, judicial clerkships, number of job offers, number of bar exam attempts, licensed status, debt, and intention to practice. See notes in Table A2 for description of baseline controls. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A9: Correlation between Partner Promotion and Predicted Performance

Dep Var: Equity Partner	(1)
Predicted Intensity (1 SD)	0.0799***
	(0.0140)
Male Likelihood	0.221
Female Likelihood	0.119
Observations	1,780
R-squared	0.188
Baseline Controls:	Yes

Notes: Baseline characteristics include demographic controls, ability proxies, and job characteristics. Demographic controls are race and ethnicity, age, law school graduation date, geographic location at time of initial survey, and initial marital status. Ability proxies are undergraduate and law school GPAs, U.S. News' 2003 law school ranking, participation in general law review, judicial clerkships, initial number of job offers, and bar exam attempts. Job characteristics are initial firm's size, initial hours worked, and area of law. *** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

Table A10: Estimating structural break in parent-share growth

Outcome: Change in share of parents		
Post-JD year index	Estimates	R2
1	0.133**	0.9627
	(0.0555)	
2	0.008	0.9475
	(0.0663)	
3	-0.0421	0.9491
	(0.0622)	
4	-0.0933	0.9567
	(0.0540)	
5	-0.0724	0.9531
	(0.0547)	
6	-0.121**	0.9629
	(0.0500)	
7	-0.0499	0.9498
	(0.0618)	
8	0.0188	0.9477
	(0.0662)	
9	0.0793	0.9529
	(0.0624)	
10	0.100	0.9568
	(0.0577)	

Source: AJD restricted data.
Notes: N = 20. Regression sample runs from 4 years before JD to 15 years after JD. Estimates are β_1 from the following regression model: $D_t = \beta_0 + \beta_1 \cdot 1\{T = t\} + \tau + \varepsilon_t$ where t denotes number of years relative to the year of law school graduation, D_t is the change in share of parents from $t-1$ to t, and τ is a quartic time trend. *** p< 0.01, ** p< 0.05, * p< 0.1

3 Theoretical Model

In this section, we present a theoretical framework to illustrate how a firm's negative reaction to childbirth may cause women to endogenously respond by delaying their first child. First, we discuss the conceptual framework. Then, we present our theoretical model.

3.1 Conceptual Framework

What are the predicted effects of career concerns[2] on fertility timing, and how do they differ by gender? To answer this question, we start with the firm's decision of whether to invest in the worker's human capital. Because the firm's investment decision (promotion decision) is profitable only if the worker is high-ability, the firm will seek to promote only high-ability workers.[3] In a world with perfect information, this is an easy problem to solve. The firm observes the workers' ability levels and knows which ones to promote. If the firm cannot observe the worker's ability level, however, the promotion decision becomes difficult. A key feature of these career concern models is that firms observe a signal of the worker's ability and update their beliefs accordingly. The worker is concerned about his current performance on future compensation. The firm is concerned about moral hazard, where low-ability workers masquerade as high-ability workers.[4]

An additional fact complicates this scenario. As raising children takes time and effort, parents have less to spend at work, thereby making less money for the firm.[5] A gender difference in the child-rearing cost results in mothers making less money for the firm than fathers, and this result occurs even if men and women are equal in terms of their ability and value of children. As investment is costly, the firm acts to minimize its losses from providing firm-specific training to (promoting) women and parents (Barron et al., 1993). It sets the highest promotion threshold for mothers, then for fathers, and the lowest threshold for childless workers.[6]

[2]Career concerns refer to the worker's uncertainty about his promotion prospects.

[3]This assumption is necessary to make the question interesting. If it were profitable to invest in everyone, then the firm would promote everyone. Likewise, if it were unprofitable to promote everyone, then the firm would invest in no one.

[4]There are different types of contracts (explicit or implicit) that solve this problem (Fama, 1980; Holmström, 1999). However, the focus of this paper is on the fertility decision and not the worker's effort-level. Therefore, we simply assume a linear contract where firms and workers share the profit and firms bear the full cost of investment (promotion). This will lead to a separating equilibrium, where high-types will exert more effort to distinguish themselves from the low-types (Akerlof, 1976; Landers et al., 1996; Spence, 1973, 2002).

[5]Assume that ability and effort are directly related.

[6]Another method of discrimination is to assign less-productive tasks to women and parents (Lehmann, 2013).

How do workers respond to this statistical discrimination? Lang and Manove (2011) provide some answers. They find that in the face of statistical discrimination on race, blacks over-invest in education relative to whites in an attempt to send a stronger signal. This example can easily be applied to our case of statistical discrimination on the basis of parental status and gender. Like race, gender is not manipulable. But parental status is. Therefore, workers have an incentive to delay their fertility so that they are not subject to the higher promotion threshold for parents. Moreover, this perverse incentive is stronger for women than for men.

3.2 Theoretical Model

We present a two-period signalling model to illustrate the roles of an information asymmetry and statistical discrimination in the emergence of a gender fertility gap.[7] Our model is built around Thomas (2015), but differs by allowing workers to have two choice variables in the first period: effort and fertility. In comparison with traditional models of career concerns that focus on effort, our focus is on the worker's fertility decision.

The key assumptions of the model are: (1) an information asymmetry (worker has private information about his marginal cost of exerting effort), (2) a gender difference in the technology of having and raising children (women bear greater costs of exerting effort than men after having a child), and (3) a productivity shock that follows a strictly-concave distribution. Under these assumptions, the firm has an incentive to hold mothers and parents to a higher standard when determining promotions. This results in a signalling game where women and associates with low marginal costs of exerting effort delay their child-birth until after the promotion decision. The model makes clear that the trade-off between increasing effort and having a child increases the signalling incentive, which is most pronounced for females with low marginal costs of effort.

3.2.1 Model setup

We present a two-period model with no discounting. There are two types of agents: firms and workers, and both are assumed to be risk-neutral. Firms are homogenous, but workers

[7]We use empirical facts from the data to inform the assumptions made in our model. Because most lawyers are married at the start of the survey and because we do not have yearly data on marital status, we abstract away from the marriage decision. Because very few lawyers are parents during law school, we focus on lawyers' fertility choices over their *career trajectory*, or the years after they graduate from law school. These simplifications allow us to focus on and highlight the effect of career concerns on the fertility decision without compromising the main results. We assume in our model that everyone enters and remains at private law firms, which is not an outlandish assumption as the majority of lawyers work at private law firms right out of law school. However, exit from law firms is something that our model does not feature. If we model an "up-or-out" contract to account for attrition, our core results do not change.

differ in their gender, $g \in \{F, M\}$, and their marginal cost of exerting effort, $\theta \in \{\theta_L, \theta_H\}$, which is known to themselves but unobserved by the firm. We denote workers with θ_L as "low effort-cost workers" and those with θ_H as "high effort-cost workers".[8] The two types are distributed with probabilities $Pr(\theta_L)$ and $Pr(\theta_H)$, respectively, and the distribution from which these types are drawn is common knowledge. Worker's gender and marginal cost of exerting effort are independently distributed.

The timing of the model is as follows:

1. At the beginning of Period 1, the worker chooses how much effort to exert and whether or not to have a child.

2. Worker receives a shock to his Period-1 output. Workers do not know their productivity draws at the time of their effort and fertility decisions.

3. At the end of Period 1, the firm observes the worker's output, which is a noisy signal of the worker's effort choice, and decides whether or not to promote the worker.

4. At the beginning of Period 2 and after receiving investment (if promoted), the worker decides how much effort to exert.

Next, we describe the firm's and worker's payoffs in turn. The firm has utility over its promotion decision $\tau \in \{0, 1\}$, which takes place at the end of Period 1 and costs c_τ, and the worker's equilibrium level of effort (e^*). Its per-period utility is as follows:

$$\Pi(\tau; e^*) = (1 - S)\cdot y - c_\tau = (1 - S)\cdot \Big(\alpha(\tau)\cdot e^*\Big) - c_\tau \tag{2}$$

where $(1 - S)$ is the firm's pre-determined share of rents and y is the worker's output. $\alpha(\tau)$ denotes the worker's human-capital, which is a function of the firm's investment decision. We assume $\alpha(1) > \alpha(0)$.

In Period 1, workers receive a one-time shock, υ, to their Period-1 output, y_1, after choosing a level of effort. Firms, however, only observe the workers' output, which is a noisy signal of their choice of effort:

$$y_1 = \alpha(0)\cdot (e_1 + \upsilon), \quad \upsilon \sim F \tag{3}$$

where F is a strictly concave distribution. Although a common assumption is a normally-distributed productivity shock, our sample of lawyers will draw from the right of the popu-

[8] These can also be thought of as an ability measure where a worker with higher ability exerts less effort to produce one unit of output relative to a worker with less ability.

lation distribution. As such, a strictly concave distribution is a logical assumption.In Period 1, no worker is promoted and so $\tau = 0$.

The salary contract for each period is pre-determined such that firms and workers share the rent to human capital. Thus, the firm receives $(1 - S_1)$ in Period 1 and $(1 - S_2)$ in Period 2, and the worker receives the remaining share. The salary contract is tied to the worker's output, which makes sense in the context of lawyers receiving year-end bonuses.[9]

The firm bears a cost to promoting the worker, which involves an investment in the worker's human capital: c_τ, with $c_1 > 0$ and $c_0 = 0$. This is a realistic assumption in the case of law firms. Firms are willing to pay the investment cost because the information gathered by them about workers' effort-costs is kept internal (not transmitted to outsiders) and so firms earn rents on trained workers (Acemoglu and Pischke, 1998). The important point is that workers are unable to fully internalize the investment cost, and thus the firm forms beliefs about workers' effort costs using their first-period signal.

We assume the cost of investment is such that it is beneficial for the firm to promote a worker with low marginal cost of effort (θ_L), but costly otherwise. The other two cases are trivial and uninteresting. If the cost of promotion were lower than its benefit, then the firm would invest in every worker regardless of his effort-cost. Similarly, if the cost of promotion were greater than its benefit, then the firm would invest in no one.

All workers participate in the labor market for two periods. Workers have utility over their effort-level (e) and whether or not they have a child (n). Their per-period utility is as follows:

$$\text{Period 1:} \qquad S_1 \cdot y_1 + \phi(n) - \frac{1}{2}(\theta + \lambda_g n) \cdot e_1^2 \tag{4}$$

$$\text{Period 2:} \qquad S_2 \cdot y_2 + \phi(n) - \frac{1}{2}(\theta + \lambda_g n) \cdot e_2^2 \tag{5}$$

where $y_1 = \alpha(0) \cdot (e_1 + \upsilon)$ and $y_2 = \alpha(\tau) \cdot e_2$. The first term is the worker's work-related utility, the second term is the net benefit from having a child, and the third term is the disutility of work.

In Period 1, workers choose an effort-level and whether or not to have a child. In Period 2, workers choose only their effort-level. In Section 3.4 we relax this assumption and allow the worker to also decide whether or not to have a child in Period 2. The comparative statics

[9] Alternatively, we can specify a salary contract where the associate worker is initially paid a fixed salary and partners equally share profits (as in Landers et al. (1996)). That will worsen the moral hazard problem, but it will not change our results. Females will still have an incentive to signal to the firm by delaying their fertility. Our results are pretty robust to the type of salary contract, as long as there is an information asymmetry, future compensation is tied to a measure of current performance, and productivity follows a strictly concave distribution.

do not change. The main difference is that firms will have higher promotion thresholds for all females instead of just mothers because of the possibility that a woman with no child in Period 1 can have one in Period 2 and reduce her effort-level.

There is a fixed benefit to having a child as well as a fixed cost. The fixed benefit can be considered the "joy" of being a parent, and this joy does not differ by gender or marginal cost of effort. The fixed cost of having a child can be thought of as a monetary cost, for example, the cost of buying diapers or purchasing baby food. The net benefit of deciding to have a child in Period 1, therefore, is denoted as $\phi(n)$.

The disutility of work is a function of the worker's effort, presence of a child, marginal cost of exerting effort, and gender: $\frac{1}{2}(\theta + \lambda_g n)\cdot e^2$. The worker's gender and fertility decision affect the disutility of work through the child, that is, through the indirect cost of having children on exerting effort, λ_g. Specifically, we assume that parents bear a greater cost of exerting effort, and conditional on having a child, females more so than men: $\lambda_F > \lambda_M > 0$.

Next, we describe the model's equilibrium thresholds and the comparative statics of interest: how does the worker's effort and fertility decisions depend on his marginal cost of exerting effort, θ? We use monotone comparative statics to answer this question.[10]

3.2.2 Equilibrium thresholds

Proposition 1 *The firm sets different promotion thresholds by gender and parental status.*

Proof. The maximization problem of the worker in Period 2 is as follows:

$$\max_{e_2} \quad v_2(e_1, n; \theta, g, \tau) = S_2\cdot \Big(\alpha(\tau)\cdot e_2\Big) + \phi(n) - \frac{1}{2}(\theta + \lambda_g n)\cdot e_2^2 \tag{6}$$

The worker's optimal effort decision in Period 2 is

$$e_2^* = \frac{S_2\cdot \alpha(\tau)}{\theta + \lambda_g n} \tag{7}$$

Firm will choose to promote a worker ($\tau^* = 1$) if and only if the expected gain from

[10] Although there are other approaches to obtain comparative statics, such as the implicit function theorem, we choose to use monotone comparative statics for its simplicity in highlighting key predictions of the model.

promotion is greater than or equal to the cost of promotion:

$$\begin{aligned} \Pi_{\theta_L}(\tau = 1) &\geq \Pi_{\theta_L}(\tau = 0) && (8)\\ Pr(\theta_L|y_1)\cdot x(\theta_L, n, g) &\geq c_1 && (9)\\ Pr(\theta_L|y_1) &\geq A\cdot(\theta_L + \lambda_g n) \equiv \tilde{x}(n, g) && (10) \end{aligned}$$

where $\Pi(\tau) = (1 - S_2)\cdot\big(\alpha(\tau)\cdot e_2^*(\tau)\big) - c_\tau$ is the firm's share of rents,$Pr(\theta_L|y_1)$ is the firm's belief that a θ_L worker sent signal y_1, $A = \frac{c_1}{S_2\cdot(1-S_2)\cdot(\alpha(1)^2-\alpha(0)^2)}$, and $x(\theta_L, n, g) = \frac{S_2\cdot(1-S_2)\cdot\left(\alpha(1)^2-\alpha(0)^2\right)}{\theta_L+\lambda_g n}$ is firm's gain from promoting worker.

The firm's net gains to promoting a worker with marginal cost of effort θ_L increases in n and in g:

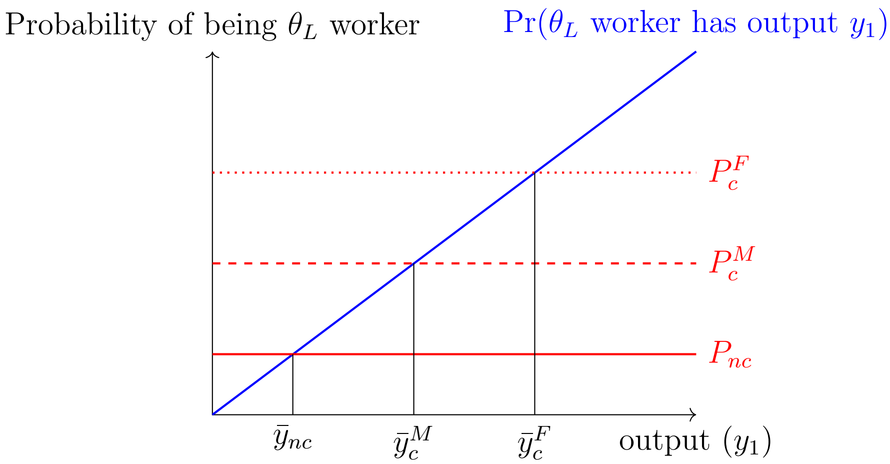
$$\begin{aligned} \frac{\partial\tilde{x}(n,g)}{\partial n} &= \frac{\partial\left(\frac{S_2\cdot(1-S_2)\cdot\left(\alpha(1)^2-\alpha(0)^2\right)}{\theta_L+\lambda_g n}\right)}{\partial n} > 0 && (11)\\ \frac{\partial\tilde{x}(n,g)}{\partial g} &= \frac{\partial\left(\frac{S_2\cdot(1-S_2)\cdot\left(\alpha(1)^2-\alpha(0)^2\right)}{\theta_L+\lambda_g n}\right)}{\partial\lambda_g} > 0 && (12) \end{aligned}$$

■

This can be solved by backwards induction. First, workers with high effort-costs work less in Period 2 than those with low effort-costs. More importantly, those with children work less than those with no children, and there is no gender difference in optimal effort-levels among those with no children (conditional on everything else). Conditional on having children, however, women exert less effort than men due to the greater indirect cost of having children.

This means that the firm earns less from workers with children and, conditional on being a parent, less from females than from males. This implies that the signal cutoff for promotion is highest for females with children and lowest for those with no children. Figure A6 depicts this result. The blue line represents the likelihood of a low effort-cost worker sending a signal y_1. This likelihood is increasing in y_1 because workers with low marginal costs of effort have higher Period-1 equilibrium levels of effort than workers with high marginal costs of effort. The horizontal red lines represent the minimum likelihood that the firm is willing to accept in order to promote a worker. Specifically, the firm is willing to accept a lower likelihood for a worker with no children than it is willing to accept for a worker with children $((P_{nc} < P_c^M < P_c^F)$. These likelihoods are inversely related to the firm's net gain from promotion. As mentioned before, the net gain to the firm from promotion is highest for workers with no children and lowest for women with children. Therefore, the inverse of the net gain is smallest for workers with no children and highest for women with children.

Figure A6: Firm's promotion thresholds



Notes: This figure shows the firm's promotion thresholds for a childless worker ($\bar{y}_{nc}$), a male worker with a child ($\bar{y}_c^M$), and a female worker with a child ($\bar{y}_c^F$). The promotion thresholds are the intersection of the blue line, which depicts the probability that a worker with output y_1 has a low marginal cost of effort, and the horizontal red lines, which represent the minimum likelihoods that the firm is willing to accept in order to promote a worker.

The intersection of the blue and red lines depicts the signal cutoff or the promotion threshold. Workers with a signal above $\bar{y}$ will be promoted while those with a signal below $\bar{y}$ will not be promoted. Because the worker's gender and parental status are observable, the firm sets different signal cutoffs for workers with children and by gender. Specifically, the firm has the highest signal cutoff for women with children and the lowest signal cutoff for workers with no children ($\bar{y}_{nc} < \bar{y}_c^M < \bar{y}_n^F$).

An alternative solution is for the firm to set gender-specific wage schemes to account for the fact that women are more likely to reduce hours or leave the labor force during child-rearing years. However, that sort of blatant discrimination is illegal. One reason that firms are able to set different promotion rates by gender and parental status without prosecution is that ability-levels are difficult to observe. As Lazear and Rosen (1990) point out, it is more difficult to document a pattern of higher promotion rates for men of a given ability relative to women than it is to document that wages in a given job differ systematically by gender.

Proposition 2 *The gain in utility from increased effort is decreasing in the marginal cost of effort.*

Proof. The worker's maximization problem in Period 1 is[11]:

$$\max_{e_1,n} \quad v_1(\theta,g,\upsilon) = S_1\cdot \alpha(0)\cdot (e_1+\upsilon) + 2\phi(n) - \frac{1}{2}(\theta+\lambda_g n)\cdot e_1^2 + \frac{1}{2}\left\{\frac{S_2^2\cdot \alpha(1)^2}{\theta+\lambda_g n} - \frac{S_2^2\cdot F(\bar{y}(n,g)-e_1^*)}{\theta+\lambda_g n}\left(\alpha(1)^2-\alpha(0)^2\right)\right\} \tag{13}$$

Assume $e_1 > e_1'$. The gain from increasing effort, $v_1(e_1,n,\theta,g,\upsilon) - v_1(e_1',n,\theta,g,\upsilon)$, is

$$S_1\cdot \alpha(0)\cdot (e_1-e_1') - \frac{1}{2}(\theta+\lambda_g n)\cdot (e_1^2 - e_1'^2) + \frac{K_1\cdot \left[F\left(\bar{y}(n,g)-e_1'\right) - F\left(\bar{y}(n,g)-e_1\right)\right]}{\theta+\lambda_g n} \tag{14}$$

where $K_1 = \frac{S_2^2\cdot\left(\alpha(1)^2-\alpha(0)^2\right)}{2}$. ■

Increasing effort in Period 1 results in three changes: (1) an increase in earnings since effort is tied to wages, (2) an increase in the disutility of work since effort is costly, and (3) an increase in promotion probability since increased effort leads to a higher signal. However, the third change - an increase in promotion probability - is smaller in magnitude for workers with a high marginal cost of effort relative to workers with a low marginal cost of effort. This is because these high effort-cost workers exert less effort in general. Moreover, they have a greater disutility of work than their low effort-cost peers. Thus, on the whole, the utility gain from increased effort in Period 1 is decreasing in the marginal cost of effort.

Proposition 3 *The gain in utility from having a child is increasing in the marginal cost of effort.*

Proof. From equation (13), the gain from having a child, $v_1(e_1,1,\theta,g,\upsilon) - v_1(e_1,0,\theta,g,\upsilon)$, is given by the following expression:

$$2\phi(1) - \frac{1}{2}\lambda_g\cdot e_1^2 + \frac{K_0 - K_1\cdot F(\bar{y}(1,g)-e_1)}{\theta+\lambda_g} - \frac{K_0 - K_1\cdot F(\bar{y}(0,g)-e_1)}{\theta} \tag{15}$$

where $K_0 = \frac{S_2^2\cdot\alpha(1)^2}{2}$ and $K_1 = \frac{S_2^2\cdot\left(\alpha(1)^2-\alpha(0)^2\right)}{2}$. As we are interested in determining the sign of how this gain changes as θ increases, we can ignore the first two terms. After some rearranging, we have

$$\frac{K_1}{\theta+\lambda_g}\left[F\left(\bar{y}(0,g)-e_1\right) - F\left(\bar{y}(1,g)-e_1\right)\right] - \frac{\lambda_g}{\theta^2+\theta\lambda_g}\left[K_0 - K_1\cdot F\left(\bar{y}(0,g)-e_1\right)\right] \tag{16}$$

[11] See Section 3.3 for derivation.

The first term reflects the indirect effect that the child has on the probability of promotion. It is negative because the firm increases the promotion threshold for workers with children as they are expected to exert less effort in Period 2 relative to those without children. The second term reflects the fact that having children also has an impact on future earnings. This term is also negative. Thus, it is straightforward to see that an increase in the marginal cost of exerting effort, θ, lowers the cost of having children. ■

The key mechanism for this proposition is that having a child in Period 1, which is before the promotion decision, subjects one to a higher promotion threshold. In addition, workers with a high marginal cost of exerting effort have a lower probability of being promoted relative to their low effort-cost peers. Because high effort-cost workers are less likely to be promoted anyway, the negative effect that a child in Period 1 has on the promotion probability is smaller relative to the negative effect for workers with low marginal costs.

Proposition 4 *Increasing effort and having a child are substitutes.*

Proof. The following expression shows the gain associated with increasing effort (from equation (16)) that depends on parental status:

$$-\frac{1}{2}\lambda_g \cdot (e_1^2 - e_1'^2) - \lambda_g \frac{K_1 \cdot \Big[F\left(\bar{y}(n,g) - e_1'\right) - F\left(\bar{y}(n,g) - e_1\right)\Big]}{(\theta + \lambda_g n)^2} \tag{17}$$

Increasing effort increases the disutility of work. This disutility is more negative for those with children. This is the first term in the equation above. Increasing effort also leads to a higher promotion probability. The second term denotes the impact of having a child on this higher promotion probability due to increased Period-1 effort. For those with children, the marginal gain is smaller relative to the marginal gain for those without children. Formally,

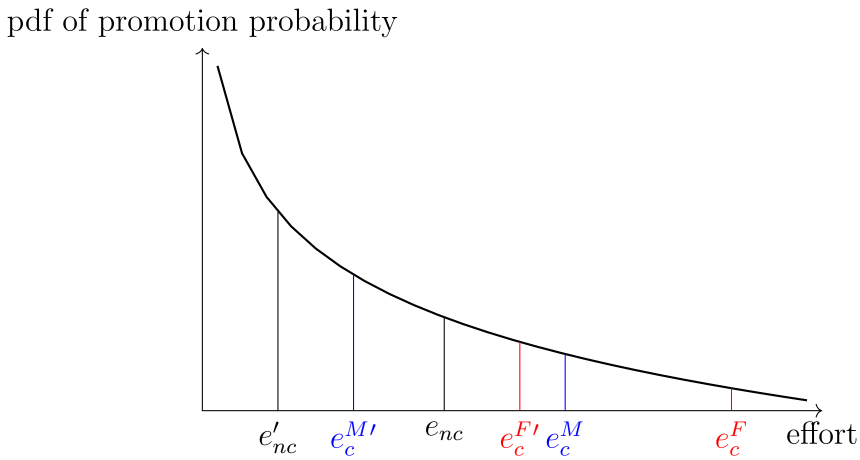
$$F(\bar{y}(0,g) - e_1') - F(\bar{y}(0,g) - e_1) > F(\bar{y}(1,g) - e_1') - F(\bar{y}(1,g) - e_1) \tag{18}$$

This inequality is true because F is strictly concave. ■

The key insight here is that children have a negative effect on the probability of promotion (through the increased promotion thresholds) but increasing effort has a *positive* effect on the probability. Figure A7 provides a graphical representation. The region under the curve and between the black vertical lines denotes the increase in promotion probability due to increased Period-1 effort for those without children. Increasing effort (moving from e_{nc}' to e_{nc}) makes it more likely that the worker will send a signal above the cutoff. If the worker is male and has a child in Period 1, the increase in promotion probability from increased effort is smaller, represented by the area under the curve and between the blue vertical lines ($e_c^{M\prime}$

to e_c^M). This can be seen visually in Figure A7 as $F(\upsilon)$ is strictly concave. Intuitively, a male worker with a child is subject to a higher promotion threshold than childless workers. Therefore, he must exert greater effort to be promoted. As firms have a higher promotion threshold for those with children relative to those without children, having a child offsets the increase in promotion probability due to increased effort through its negative impact on the odds of promotion. For females with a child, this offset will be even greater, illustrated by the area under the curve and between the red vertical lines.

Figure A7: Effect of children on increased benefits of effort



Notes: This figure shows that the utility gain from an increase in effort is smaller for workers with a child. Specifically, it is smallest for female workers with a child, then for male workers with a child, all relative to a childless worker's utility gain.

Below are the model's main empirical predictions about fertility for our sample of lawyers:

1. Fewer female lawyers than male lawyers will have children before the partnership decision.

2. The gender difference in fertility timing (EP1) will be larger among lawyers with high effort-levels relative to the gender difference among those with low effort-levels.

3.3 Worker's Period-1 Effort and Fertility Decisions

The worker's objective function is his utility in Period 1 and his expected utility in Period 2. The Period 1 utility is as follows:

$$U_1 = S_1 \cdot \alpha(0) \cdot (e_1 + \upsilon) + \phi(n) - \frac{1}{2}(\theta + \lambda_g n) \cdot e_1^2 \tag{19}$$

His expected utility in Period 2 depends on his chances of receiving a promotion. What is the probability of promotion? It depends upon the likelihood of his signal meeting the firm's threshold.

$$
\begin{aligned}
P(\text{Promotion}) &= P(y_1 \geq \bar{y}(n,g)) \\
&= P(e_1^* + \upsilon \geq \bar{y}(n,g)) \\
&= P(\upsilon \geq \bar{y}(n,g) - e_1^*) \\
&= 1 - F(\bar{y}(n,g) - e_1^*)
\end{aligned}
\tag{20}
$$

The payoffs to the worker in Period 2 differ depending on whether or not he received a promotion:

$$
\begin{aligned}
U_2^{\text{Promotion}} &= S_2 \cdot \left(\alpha(1) \cdot e^*_{2,\tau=1}\right) + \phi(n) - \frac{1}{2}(\theta + \lambda_g n) \cdot e^{*2}_{2,\tau=1} \\
&= S_2 \cdot \left(\alpha(1) \cdot \left(\frac{S_2 \cdot \alpha(1)}{\theta + \lambda_g n}\right)\right) + \phi(n) - \frac{1}{2}(\theta + \lambda_g n) \cdot \left(\frac{S_2 \cdot \alpha(1)}{\theta + \lambda_g n}\right)^2 \\
&= \frac{S_2^2 \cdot \alpha(1)^2}{\theta + \lambda_g n} + \phi(n) - \frac{1}{2}(\theta + \lambda_g n) \cdot \frac{S_2^2 \cdot \alpha(1)^2}{(\theta + \lambda_g n)^2} \\
&= \frac{1}{2} \cdot \frac{S_2^2 \cdot \alpha(1)^2}{\theta + \lambda_g n} + \phi(n)
\end{aligned}
\tag{21}
$$

$$
\begin{aligned}
U_2^{\text{No Promotion}} &= S_2 \cdot \left(\alpha(0) \cdot e^*_{2,\tau=0}\right) + \phi(n) - \frac{1}{2}(\theta + \lambda_g n) \cdot e^{*2}_{2,\tau=0} \\
&= S_2 \cdot \alpha(0) \cdot \left(\frac{S_2 \cdot \alpha(0)}{\theta + \lambda_g n}\right) + \phi(n) - \frac{1}{2}(\theta + \lambda_g n) \cdot \left(\frac{S_2 \cdot \alpha(0)}{\theta + \lambda_g n}\right)^2 \\
&= \frac{S_2^2 \cdot \alpha(0)^2}{\theta + \lambda_g n} + \phi(n) - \frac{1}{2}(\theta + \lambda_g n) \cdot \frac{S_2^2 \cdot \alpha(0)^2}{(\theta + \lambda_g n)^2} \\
&= \frac{1}{2} \cdot \frac{S_2^2 \cdot \alpha(0)^2}{\theta + \lambda_g n} + \phi(n)
\end{aligned}
\tag{22}
$$

The worker's expected utility in Period 2 is as follows:

$$
\begin{aligned}
\mathbb{E}(U_2) &= P(\text{Promotion}) \cdot U_2^{\text{Promotion}} + P(\text{No Promotion}) \cdot U_2^{\text{No Promotion}} \\
&= \big(1 - F(\bar{s}(n,g) - e_1^*)\big) \cdot \left(\frac{1}{2} \cdot \frac{S_2^2 \cdot \alpha(1)^2}{\theta + \lambda_g n} + \phi(n)\right) \\
&\quad + F(\bar{s}(n,g) - e_1^*) \cdot \left(\frac{1}{2} \cdot \frac{S_2^2 \cdot \alpha(0)^2}{\theta + \lambda_g n} + \phi(n)\right) \\
&= \frac{1}{2}\left\{\frac{S_2^2 \cdot \alpha(1)^2}{\theta + \lambda_g n} - \frac{S_2^2 \cdot F(\bar{y}(n,g) - e_1^*)}{\theta + \lambda_g n} \cdot \left(\alpha(1)^2 - \alpha(0)^2\right)\right\} + \phi(n)
\end{aligned}
\tag{23}
$$

Putting these together, we have that the Period 1 maximization problem of the worker's expected utility is:

$$\max_{e_1,n} \quad v_1(\theta,g,\upsilon) = S_1\cdot\alpha(0)\cdot(e_1+\upsilon) + 2\phi(n) - \frac{1}{2}(\theta+\lambda_g n)\cdot e_1^2 + \frac{1}{2}\left\{\frac{S_2^2\cdot\alpha(1)^2}{\theta+\lambda_g n} - \frac{S_2^2\cdot F(\bar{y}(n,g)-e_1^*)}{\theta+\lambda_g n}\left(\alpha(1)^2-\alpha(0)^2\right)\right\} \tag{24}$$

3.4 Allowing worker to have child in Period 2

We now consider the worker's optimal decisions if we allow the worker to have a child in Period 2. In Period 2, the signalling game is over, and we can think of the worker's maximization problem as a simple utility maximization problem over e_2 and n subject to a budget constraint. Therefore, the worker's decision to have a child in Period 2 will depend on the relationship between the net benefit of having a child ($\phi(n)$) and the cost of having a child (in terms of increased consumption and increased disutility of work).

If ϕ is high enough such that some workers will want to choose to have a child in Period 2, the firm's promotion decision will change. Specifically, the firm will have higher promotion thresholds for childless workers, relative to the equilibrium threshold in our base case, because it is possible for workers to have children in Period 2 and they will work less. Again, because females bear a higher child-cost on effort, the promotion thresholds will differ by gender *even for childless workers.* The comparative statics will remain the same in that female workers will continue to have a greater incentive to signal to the firm by delaying child-birth.

If ϕ is not high enough such that it is not optimal for any worker to have a child in Period 2, then we are back to our base case and nothing changes.

4 Factor Analysis

4.1 Factor Score Regression

This section describes the factor indeterminancy problem that arises when using the two-step process of Factor Score Regression (FSR). In the first step, the scores on the latent variables are predicted using a factor analysis. These scores are referred to as factor scores. In the second step, the factor scores are used in an OLS regression as explanatory variables.

Devlieger, Mayer, and Rossel (2016) show that in a model with a latent independent variable and an observed dependent variable, constructing factor scores using the Regression method in the first-step will yield an unbiased coefficient estimate in the second-step. For the sake of the reader, we reproduce their proof here:

Say the structural equation is:

$$Y = \delta\varphi + \varepsilon \tag{25}$$

As φ is latent, the following measurement model is used:

$$q = A_q\varphi + \upsilon \tag{26}$$

where $q = (Q_1, \cdots, Q_i, \cdots, Q_k)^T$ are vectors of mean-centered proxy variables measuring φ, A_q is a vector of the factor loadings and υ is a vector of measurement error variables.

In the first step of FSR, we use this measurement system to calculate factor scores for φ:

$$F_\varphi = \Lambda_\varphi q \tag{27}$$

where $F_\varphi = (F_1, \cdots, F_j)$ for j total factors. The factor scores are calculated by multiplying a factor score matrix Λ_φ with proxy variables q. We fix the metric scales of φ by fixing one factor loading per latent variable to 1 (also known as unstandardized parameterization). The computation of Λ_φ depends on the method used for the prediction of the factor score. There is a variety of possible methods; Grice (2001) discusses several options.

In the second step of FSR, a linear regression is performed between the factor scores, resulting in a regression coefficient. In a simple linear regression, the true regression coefficient is defined as follows:

$$\delta = \frac{cov(\varphi, Y)}{var(\varphi)} \tag{28}$$

When performing the linear regression with factor scores, the regression coefficient becomes:

$$\beta = \frac{cov(F_\varphi, Y)}{var(F_\varphi)} \tag{29}$$

which is not necessarily the same as the true regression coefficient, δ. The relationship between δ and β is depicted as follows:

$$\begin{aligned}
\beta &= \frac{cov(F_\varphi, Y)}{var(F_\varphi)} && (30)\\
&= \frac{cov(\Lambda_\varphi q, Y)}{\Lambda_\varphi \Sigma_q \Lambda'_\varphi} && (31)\\
&= \frac{\Lambda_\varphi cov(q, Y)}{\Lambda_\varphi \Sigma_q \Lambda'_\varphi} && (32)\\
&= \frac{\Lambda_\varphi cov(A_q\varphi + \upsilon, Y)}{\Lambda_\varphi \Sigma_q \Lambda'_\varphi} && (33)\\
&= \frac{\Lambda_\varphi A_q cov(\varphi, Y)}{\Lambda_\varphi \Sigma_q \Lambda'_\varphi} && (34)\\
&= \frac{\Lambda_\varphi A_q cov(\varphi, \delta\varphi + \varepsilon)}{\Lambda_\varphi \Sigma_q \Lambda'_\varphi} && (35)\\
&= \frac{\Lambda_\varphi A_q var(\varphi)}{\Lambda_\varphi \Sigma_q \Lambda'_\varphi}\delta && (36)
\end{aligned}$$

This equation makes clear that the estimated regression coefficient β does not necessarily equal δ.

Now we discuss the Regression method of constructing factor scores used to predict Λ_φ. The Regression method gives the following estimate for the factor score matrix:

$$\Lambda^R_\varphi = \Phi A'_q \Sigma_q^{-1} = var(\varphi) A'_q \Sigma_q^{-1} \qquad (37)$$

This means that $var(F_\varphi)$ can be simplified:

$$\begin{aligned}
var(F_\varphi) &= \Lambda^R_\varphi \Sigma_q \Lambda^{R'}_\varphi && (38)\\
&= \left(var(\varphi) A'_q \Sigma_q^{-1}\right) \Sigma_q \Lambda^{R'}_\varphi && (39)\\
&= var(\varphi) A'_q I \Lambda^{R'}_\varphi && (40)\\
&= \Phi' \Lambda' \Sigma_q^{-1} \Lambda \Phi && (41)\\
&= \Lambda^R_\varphi A_q var(\varphi) && (42)
\end{aligned}$$

Plugging this into the denominator of equation (30) yields (the numerator does not change):

$$\beta = \frac{cov(F_\varphi, Y)}{var(F_\varphi)} \tag{43}$$

$$= \frac{\Lambda_\varphi A_q var(\varphi)}{\Lambda_\varphi^R A_q var(\varphi)}\delta \tag{44}$$

$$= \delta \tag{45}$$

4.2 Exploratory Factor Analysis

This section describes the empirical methodology for our factor analysis. First, we impute missing item responses using gender-specific averages to minimize the number of dropped respondents due to missing reponse. To minimize bias, we impute responses only for those with one or two missing responses in each question set. There are three question sets (Question 38, Question 40, and Question 65), each with 9-11 potential reasons that are to be ranked by the lawyer. (See Section 1.2 for exact text of the questions.) If the lawyer ranked all but one or two of the potential reasons, then we replace the missing ranking(s) with a gender-specific average ranking. We leave out the "Other" category in our factor analysis as it is too broad to systemically capture a common underlying factor. Question 40, which asks about the important determinants in choosing between multiple job offers, was asked only for those who received more than one job offer. A majority of respondents received multiple job offers (64 percent). But for those with only one offer, we impute their rankings of each potential reason using gender-specific averages. As a robustness check, we use an alternative imputation method where we use individual-specific means. Our main results do not change much.

We run two specifications for the factor analysis. First, we construct scores for everyone in the sample. Second, we construct scores separately for males and females. Our factor analysis finds that there are seven factors that, cumulatively, explain all of the variance in survey responses for all lawyers. The eigenvalues are reported in Table A11. The Kaiser test says that only factors with an eigenvalue of 1.0 or greater are meaningful, which tells us to keep the first five factors. The second test is the scree test (Figure A8). This test uses a graphical method to determine which factors to keep. The criterion is to keep the factors up until the line becomes flat or flatter. According to this visual test, there are seven common factors. We chose to keep the seven factors that explain all of the variance.

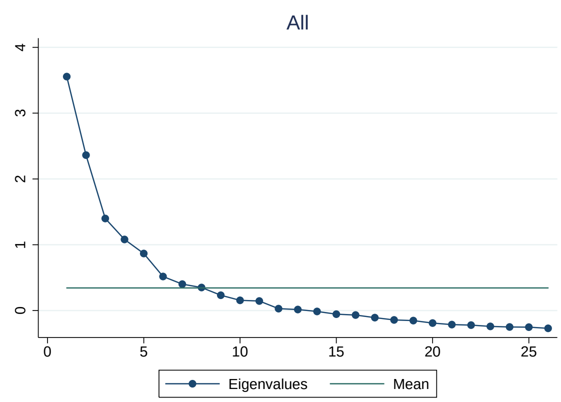
The rotated factor loadings in Table A12 report the correlation between the survey question and the factor. They provide a picture of the most important "factors" in the lawyer's career decisions, as captured by the survey questions and lawyer's responses. Factor analysis

Table A11: Principal factor analysis/correlation for everyone

Factor #	Eigenvalue	Proportion	Cumulative
1	1.71	0.19	0.19
2	1.69	0.19	0.38
3	1.43	0.16	0.54
4	1.42	0.16	0.70
5	1.00	0.11	0.81
6	0.97	0.11	0.92
7	0.89	0.10	1.02

Source: AJD restricted data.
Note: Only the seven factors with a cumulative proportion reaching 1 are shown. For full results, see the Online Appendix. LR test: independent vs. saturated: chi2(136) = 4062.42 Prob>chi2 = 0.0000

Figure A8: Scree plot



is designed to maximize the amount of variation explained by the first factor; each additional factor tries to explain as much leftover variation as possible. Looking at this table, we define the first factor as social responsibility. The remaining six are classified as: earning potential, prestige, career development, firm's ranking, mission match, and financial security.

Table A12: Rotated factor loadings and unique variances for everyone

Survey question	**Factor1**	**Factor2**	**Factor3**	**Factor4**	**Factor5**	**Factor6**	**Factor7**	**Uniqueness**
Why sector?								
Medium-to-long-term earning potential							0.5631	0.4761
Substantive interest in a specific field of law				0.5360				0.6488
Salary to pay off law school debts							0.5205	0.6295
Opportunity to develop specific skills				0.5929				0.5300
Potential to balance work and personal life				0.5213				0.5747
Opportunity to do socially responsible work	0.5063			0.5257				0.3839
Prestige of the sector			0.6920					0.4710
Opportunities for future career mobility			0.5302					0.5599
Why law school?								
Intellectual challenge of law school and the law								0.8161
Desire to help individuals as a lawyer	0.7279							0.4430
Desire to develop a satisfying career								0.6842
Desire to defer entry into the job market								0.8653
Desire for eventual financial security								0.6012
Desire to change or improve society	0.7977							0.3491
Becoming influential in a powerful profession								0.6721
Desire to build a set of transferable skills								0.7716
Why job offer?								
Salary		0.7795						0.3487
Benefits		0.7616						0.3763
Office environment/collegiality								0.5728
Hours expected								0.5163
Prospects for advancement						0.5142		0.5485
Good match of employer's mission and my own						0.5545		0.5130
Location								0.7925
Size					0.5536			0.6219
Prestige			0.4702		0.5083			0.5005
Training/mentorship opportunities								0.6213

Source: AJD restricted data.

Notes: Blanks represent loadings where the absolute value is less than 0.4.

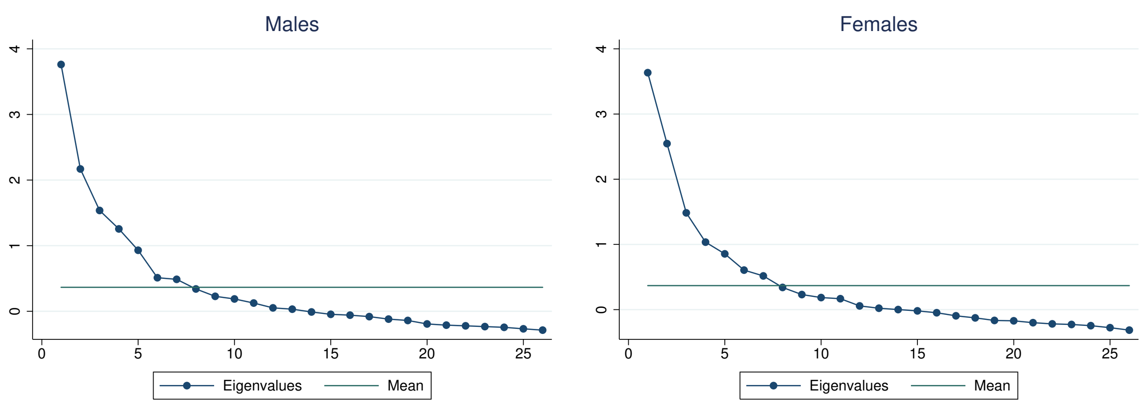
We also conduct our factor analysis separately for males and females to see if there are any gender differences in latent preferences driving career decisions. Our factor analysis again finds seven factors that explain all of the variance in survey responses. The eigenvalues are reported in Table A13. The scree test also confirms that there are seven factors.

Table A13: Principal factor analysis/correlation by gender

	Males			Females		
Factor #	**Eigenvalue**	**Proportion**	**Cumulative**	**Eigenvalue**	**Proportion**	**Cumulative**
1	1.79	0.19	0.19	1.80	0.20	0.20
2	1.69	0.18	0.37	1.63	0.18	0.38
3	1.62	0.17	0.54	1.63	0.18	0.56
4	1.50	0.16	0.70	1.62	0.18	0.74
5	1.06	0.11	0.81	1.44	0.16	0.90
6	1.05	0.11	0.93	0.90	0.10	1.00
7	0.89	0.09	1.02	0.83	0.09	1.09

Source: AJD restricted data.
Note: Only the seven factors with a cumulative proportion reaching 1 are shown. For full results, see the Online Appendix. LR test: independent vs. saturated: chi2(136) = 4062.42 Prob>chi2 = 0.0000

Figure A9: Scree plot



Tables A14 and A15 report the rotated factor loadings for males and females, respectively. Perhaps not surprisingly, the first and most important factor differs by gender. Specifically, males care most about earning potential in making their career decisions while females care most about the office environment and "fit". The factors for men are: earning potential, social responsibility, prestige, career development, financial security, office environment, and career goals. The factors for women are: office environment and "fit", earning potential, social responsibility, prestige, career development, financial security, and stability.

Table A14: Rotated factor loadings and unique variances for males

Survey question	**Factor1**	**Factor2**	**Factor3**	**Factor4**	**Factor5**	**Factor6**	**Factor7**	**Uniqueness**
Why sector?								
Medium-to-long-term earning potential					0.5891			0.4740
Substantive interest in a specific field of law				0.5736				0.5950
Salary to pay off law school debts					0.5356			0.6387
Opportunity to develop specific skills				0.6434				0.5047
Potential to balance work and personal life				0.4550		0.4528		0.5465
Opportunity to do socially responsible work		0.5490		0.4539				0.3692
Prestige of the sector							0.5580	0.4926
Opportunities for future career mobility							0.4878	0.5205
Why law school?								
Intellectual challenge of law school and the law								0.7810
Desire to help individuals as a lawyer		0.6916						0.4913
Desire to develop a satisfying career								0.7055
Desire to defer entry into the job market								0.8482
Desire for eventual financial security								0.6264
Desire to change or improve society		0.7974						0.3572
Becoming influential in a powerful profession								0.6929
Desire to build a set of transferable skills								0.7864
Why job offer?								
Salary	0.8030							0.3152
Benefits	0.8073							0.3071
Office environment/collegiality						0.4604		0.5586
Hours expected						0.6193		0.4843
Prospects for advancement								0.5226
Good match of employer's mission and my own								0.4994
Location								0.7772
Size			0.6223					0.5646
Prestige			0.7093					0.4407
Training/mentorship opportunities			0.5013					0.5686

Source: AJD restricted data.

Notes: Blanks represent loadings where the absolute value is less than 0.4.

Table A15: Rotated factor loadings and unique variances for females

Survey question	**Factor1**	**Factor2**	**Factor3**	**Factor4**	**Factor5**	**Factor6**	**Factor7**	**Uniqueness**
Why sector?								
Medium-to-long-term earning potential						0.5792		0.4704
Substantive interest in a specific field of law					0.5153			0.6672
Salary to pay off law school debts						0.5164		0.5976
Opportunity to develop specific skills					0.5728			0.5345
Potential to balance work and personal life					0.5477			0.5816
Opportunity to do socially responsible work			0.4336		0.5836			0.3974
Prestige of the sector				0.7470				0.4143
Opportunities for future career mobility				0.5097				0.5662
Why law school?								
Intellectual challenge of law school and the law								0.8353
Desire to help individuals as a lawyer			0.7467					0.4160
Desire to develop a satisfying career								0.6510
Desire to defer entry into the job market								0.8484
Desire for eventual financial security							0.4986	0.5480
Desire to change or improve society			0.7898					0.3483
Becoming influential in a powerful profession								0.6356
Desire to build a set of transferable skills							0.4700	0.7204
Why job offer?								
Salary		0.7564						0.3770
Benefits		0.7223						0.4255
Office environment/collegiality	0.5192							0.5852
Hours expected	0.5224							0.5247
Prospects for advancement	0.5371							0.5550
Good match of employer's mission and my own	0.6168							0.5211
Location								0.7405
Size								0.6411
Prestige								0.5281
Training/mentorship opportunities	0.4987							0.6384

Source: AJD restricted data.

Notes: Blanks represent loadings where the absolute value is less than 0.4.

Factor analysis assumes that the error terms are governed by a single latent factor, and therefore uses a correlation matrix of observed variables to extract this latent factor. If the observed variables are completely non-collinear, then factor analysis would extract as many as factors as variables from the correlation matrix. That is, each observed variable would be its own factor. A good validity test, therefore, would be one that measures the degree to which the observed variables share a common factor. That is, is the correlation matrix "factorable"?

There are two tests for this. Barlett's test of sphericity calculates the determinate of the matrix, which is then converted to a chi-square statistic and tested for significance. If it is statistically significant, then we can reject the null hypothesis that the observed variables are non-collinear. The determinant of the correlation matrix is 0.001 for both males and females, providing a p-value of 0. Since this is highly statistically significant, we can proceed with factor analysis.

The Kaiser-Meyer-Olkin measure of sampling adequacy test (KMO) tests the validity of the observed variables sharing a common factor. If two variables share a common factor with other variables, their partial correlation, which indicates the unique variance they share, will be small. In particular, the KMO is calculated as follows:

$$KMO = \frac{\sum_i \sum_j r_{ij}^2}{\sum_i \sum_j r_{ij}^2 + (\sum_i \sum_j a_{ij}^2)}$$

Scores between 0.9 and 1.0 are ideal, while scores below 0.6 are "miserable" and factor analysis is not recommended. The table of KMO interpretations is below. Our values of 0.739 for males, 0.745 for females, and 0.754 for everyone is "middling" indicating that the factors extracted will account for a fair amount of variance, but not a substantial amount.

Cronbach's alpha is a rule-of-thumb rather than a statistical test, but it tells us how correlated the set of items being tested are correlated with one latent factor. The rule-of-thumb is that the coefficient should be least 0.50, with it ideally being at 0.70 or higher. The Cronbach's alpha for everyone is 0.777, for males is 0.785, and is 0.767 for females, all above our threshold of 0.70.

Table A16: Robustness checks

Robustness tests	Male	Female	Everyone
Determinant of correlation matrix	0.001	0.001	0.001
Bartlett test of sphericity (p-value)	0.000	0.000	0.000
Kaiser-Meyer-Olkin Measure of Sampling Adequacy	0.739	0.745	0.754
Cronbach's alpha	0.785	0.767	0.777

References

Acemoglu, D. and J.-S. Pischke (1998). Why Do Firms Train? Theory and Evidence. *Quarterly Journal of Economics 113*(1), 79–119.

Akerlof, G. (1976). The economics of caste and of the rat race and other woeful tales. *The Quarterly Journal of Economics*, 599–617.

Barron, J. M., D. A. Black, and M. A. Loewenstein (1993). Gender differences in training, capital, and wages. *Journal of human Resources*, 343–364.

Fama, E. F. (1980). Agency Problems and the Theory of the Firm. *Journal of Political Economy*, 288–307.

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies 66*(1), 169–182.

Landers, R. M., J. B. Rebitzer, and L. J. Taylor (1996). Rat Race Redux: Adverse selection in the determination of work hours in law firms. *American Economic Review*, 329–348.

Lang, K. and M. Manove (2011). Education and labor market discrimination. *American Economic Review 101*(4), 1467–1496.

Lazear, E. P. and S. Rosen (1990). Male-female wage differentials in job ladders. *Journal of Labor Economics*, S106–S123.

Lehmann, J.-Y. K. (2013). Job assignment and promotion under statistical discrimination: Evidence from the early careers of lawyers.

Pan, J. (2015). Gender segregation in occupations: The role of tipping and social interactions. *Journal of Labor Economics 33*(2), 365–408.

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics*, 355–374.

Spence, M. (2002). Signaling in retrospect and the informational structure of markets. *American Economic Review 92*(3), 434–459.