Appendix: For Online Publication

Appendix 1: Data

A1.1: Sample Definitions

XMZYJS:

XMZYJS gave us microdata on all the ads for jobs that appeared in calendar year 2010. To construct our analysis sample from the population of ads in our frame, we excluded ads with the following characteristics: a minimum requested age below 18 or a maximum requested age over 60; an offered wage above 10,000 yuan/month; missing number of vacancies; jobs located outside the city of Xiamen; required education level is illiterate/other/undergraduate or above; missing community information; jobs that require a *zhicheng* rank above the assistant level; jobs in agriculture, ads with missing education requirement; missing occupation; and ads with missing firm information.

XMRC:

XMRC gave us microdata on all the ads for jobs that received their first application between May 1 and October 30, 2010; our XMRC sampling criteria are almost identical to those in XMZYJS. To construct our analysis sample from the population of ads in our frame, we excluded ads with a minimum requested age below 18 or a maximum requested age over 60; ads offering more than 10,000 yuan/month or less than 1,000 yuan/month; ads requesting a master's, professional or PhD degree (all of these were rare); ads for more than 10 vacancies (since job descriptions tended to be vague); ads for jobs located outside Xiamen city; ads with missing firm information; ads from firms located outside mainland China; ads with missing occupation information; and ads without a stated education requirement.

Zhaopin:

Our Zhaopin data are the same as the data used in KS: all unique ads posted in four five-week observation periods in 2008-2010. These data were scraped from Zhaopin's site using a custom-designed web crawler that visited the site nightly, and run through a de-duplication algorithm that eliminated re-postings of the same ad. Additional details on construction of our analysis sample are available in KS. They are very similar to the procedures in the XMZYJS and XMRC data.

Computrabajo:

To construct an analysis sample from the Computrabajo website, we collected advertisements daily for approximately 18 months between early 2011 and mid-2012 using a web crawler. Both the standardized fields and the open text portions of each ad were parsed to extract variables for the analysis. We use the universe of unique advertisements posted during this period.[1] As in all the data sets described in Table 1, our Computrabajo analysis sample restricts attention to ads where the required level of education is explicitly stated. Ads offering a monthly salary below 1,700 Mexican pesos, which is below the minimum wage, or above 100,000 Mexican pesos per month were dropped from the sample.

Notably, each of our four datasets is a 'stock' sample of vacancies that were unfilled at some point during a sampling period. In datasets like XMZYJS, where the sampling period is an entire calendar year and most vacancies last only a few weeks, this leads to only a minor overrepresentation of long-duration vacancies. Still, to assess the likely effects of duration-bias in sampling in our analyses, we replicated our main analysis in the Zhaopin data —which has the shortest sampling windows—for an 'inflow' sample of ads that was first posted well after the start of our observation window. There was very little change in the results.

[1] See Delgado Helleseter (2013) for additional details on de-duplication procedures and sample construction.

A1.2: How is Beauty Requested in Zhaopin and Computrabajo ads?

Table A1.2: Ten Most Frequent Beauty Requests, Zhaopin and Computrabajo Data

	(a) Zhaopin Data		
Rank	**Percentage**	**Chinese text**	**Translation**
1	30.05	形象气质佳	good image and temperament
2	9.54	五官端正	has regular facial features
3	7.35	形象良好	good image
4	5.76	品貌端正	well-shaped figure and decorous/straight appearance
5	4.99	形象好	good image
6	4.56	形象气质	image and temperament
7	4.22	形象佳	good image
8	4.11	形象好，气质佳	good image and temperament
9	3.68	相貌端正	good appearance
10	2.49	形象气质良好	good image and temperament
Others	23.25		
Total	100.00		
	(b) Computrabajo Data		
Rank	**Percentage**	**Spanish text**	**Translation**
1	52.09	Excelente presentación	Great appearance
2	24.94	Buena presentación	Good appearance
3	7.13	Buena presencia	Good appearance (presence)
4	6.73	Excelente presencia	Great appearance (presence)
5	4.31	Buena imagen	Good image
6	2.56	Sin tatuajes	No tattoos
7	0.37	Presentación excelente	Great appearance
8	0.17	Buena apariencia	Good appearance
9	0.09	Ser talla	Be of certain (dress) size
10	0.06	Peso acorde a estatura	Weight appropriate for height
Others	1.56		
Total	100.00		

Table A1.3: Broad Occupation Distributions on Job Boards and Among Private Sector Workers

Broad Occupation	(1)	(2)	(3)	(4)	(5)	(6)
	China				Mexico	
	XMZYJS Ads	XMRC Ads	Zhaopin Ads	Private Sector Workers	Computrabajo	Private Sector Workers
Management	3.98	2.09	2.03	4.27	1.32	6.22
Sales and Procurement	13.1	16.63	22.18	21.72	23.46	20.41
Service Occupations	31.71	12.71	10.92	21.65	17.06	28.52
Professional/Technical	13.02	30.31	47.13	8.35	35.87	15.45
Production, Construction, Manufacturing	33.9	30.71	13.41	44.00	22.30	29.39
Other	4.29	7.53	4.34	0		
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Sources: XMZYJS, XMRC, Zhaopin and Computrabajo job boards, plus authors' calculations from 2005 Chinese Census (1%) and 2010 Mexican Census microdata files.

Appendix 2: Evidence on the Job-Specificity of Employers' Gender Requests

A2.1: Evidence on Gender-Targeting at the Firm Level

Since the vast majority of employers post multiple ads on any given job board, it may be of interest to ask how frequently a typical *firm* engages in gender targeting. To that end, KS examined the incidence of gender-targeting by firm (rather than by ad). One key result was that the share of firms that issued a gender-targeted ad at some point during the 20-week observation period is much higher than the share of *ads* that are gender targeted. Specifically, among firms who placed more than 50 Zhaopin job ads, 70.7 percent issued at least one gender-targeted ad during the 20-week observation window. In XMZYJS, XMRC and Computrabajo respectively, Table A2.1 shows that these shares were 98.8, 100, and 98.6 percent respectively. While these higher numbers reflect both a greater overall incidence of gender targeting in the new datasets and the longer observation windows, they illustrate that gender-targeting of job ads is extremely common in all these labor markets.

KS also found that many firms gender-target in both directions, issuing some ads that request men and others that request women. Specifically, among firms who placed more than 50 Zhaopin job ads, 38.6 percent placed at least one ad that favors men *and* one ad that favors women. In XMZYJS, XMRC, and Computrabajo respectively, Table A2.1 shows that these shares were 93.4, 96.2, and 89.2 percent respectively. This dramatically illustrates that a large share of the variation in advertised gender preferences in all these markets takes place *within firms*: models based on *firm-level* tastes for men relative to women are likely to be of little help in understanding these patterns. A more appropriate class of models would view firms as having a portfolio of jobs, some of which they view as best suited to men, and others to women.

Table A2.1: Advertised Gender Preferences, by Firm

Total number of ads placed by the firm	Share of firms requesting: Any gender	Men	Women	Both genders	*N*
(a) XMZYJS Data					
1	0.768	0.310	0.458	0.000	1,540
2-10	0.873	0.647	0.658	0.433	4,637
11-50	0.975	0.888	0.863	0.775	2,074
51 and more	0.988	0.973	0.949	0.934	588
All firms	0.886	0.667	0.691	0.471	8,839
(b) XMRC Data					
1	0.432	0.162	0.270	0.000	1,213
2-10	0.758	0.442	0.587	0.272	4,549
11-50	0.954	0.796	0.868	0.710	892
51 and more	1.000	1.000	0.962	0.962	26
All firms	0.726	0.441	0.569	0.284	6,680
(c) Zhaopin Data					
1	0.165	0.056	0.109	0.000	12,692
2-10	0.326	0.155	0.224	0.053	40,932
11-50	0.548	0.344	0.392	0.189	16,239
51 and more	0.707	0.535	0.558	0.386	3,773
All firms	0.367	0.199	0.258	0.091	73,636
(d) Computrabajo Data					
1	0.343	0.135	0.207	0.000	458
2-10	0.633	0.369	0.447	0.183	807
11-50	0.887	0.698	0.788	0.598	391
51 and more	0.986	0.920	0.958	0.892	288
All firms	0.668	0.462	0.535	0.329	1,944

A2.2: Variance Decomposition: The Role of Occupations, Firms and Jobs in Gender-Targeting

Pursuing the above idea further, KS also decomposed the variance of P^F, P^M and P^F+P^M across ads into components that are associated with occupations, with firms, with differences in the way firms gender their occupations, and within job (firm*occupation) variation. Table **A2.2** reproduces that table, and replicates it for the XMRC, XMZYJS, and Computrabajo datasets respectively. Methodological details of the decomposition are provided in KS. Common features of all four datasets include the following. First, occupation alone explains very little of the variation in advertised gender preferences (less than 10 percent in all cases and usually much less). Second, in all three Chinese datasets, occupation and firm effects together (but not interacted) explain between 31.3 and 39.1 percent of the variance in advertised gender preferences (this number is 17.8 percent in Computrabajo). Thus, in all four datasets *the majority of the variance in advertised gender preferences is within firms*: it is not so much that some firms prefer to hire women and others prefer men, but more that most firms prefer men for one subset of their jobs and women for a different subset.

Third, adding occupation*firm interactions accounts for between 9.3 and 31.2 percent of the variance. Finally, in all datasets there is a large degree of within-job-cell variation in advertised gender preferences, ranging from about a third of the total in Zhaopin and XMRC to over 70 percent in Computrabajo.

Together, Tables A2.1 and A2.2 highlight the substantial heterogeneity in employers' gender preferences inside the same firm. As noted, this phenomenon is hard to reconcile with models based on *firm*-level tastes for men versus women, and more suggestive of models where all firms have some jobs they feel as more suitable for men, and others for women.

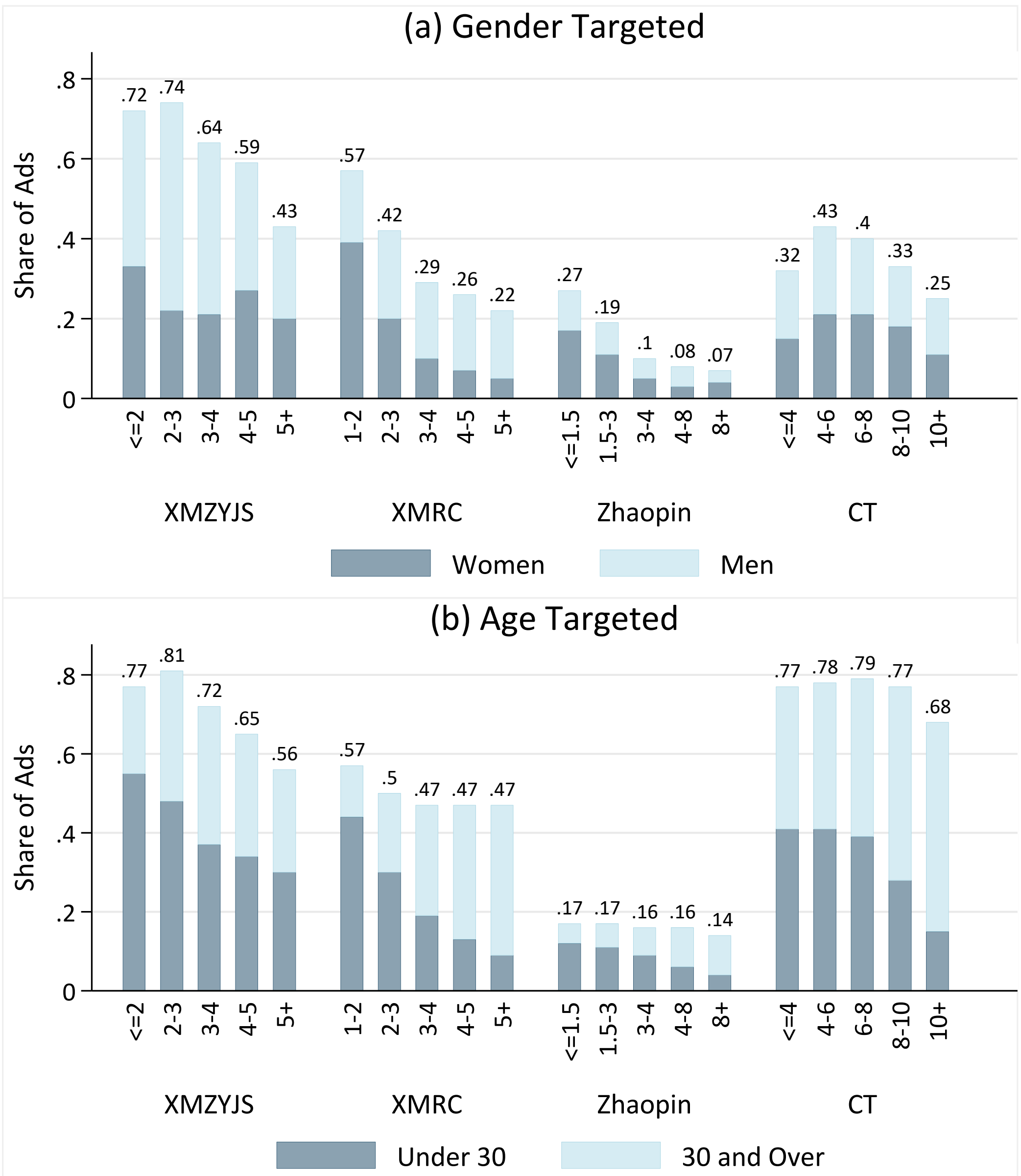
Table A2.2: Variance Decomposition

Share of variance explained by:	Dependent variable (1) Ad requests men (P^M)	(2) Ad requests women (P^F)	(3) Ad is gender-targeted (P^M+P^F)
A. XMZYJS Data			
1. Occupation	0.080	0.065	0.017
2. Firm	0.184	0.189	0.374
3. Joint occupation and firm	0.105	0.064	0.031
4. Total, occupation plus firm	0.370	0.318	0.421
5. Occupation*Firm interactions	0.142	0.159	0.190
6. Total between job cells	0.511	0.477	0.611
7. Within job cells	0.489	0.523	0.389
8. Total	1.000	1.000	1.000
B. XMRC Data			
1. Occupation	0.059	0.084	0.037
2. Firm	0.223	0.244	0.276
3. Joint occupation and firm	0.062	0.031	0.035
4. Total, occupation plus firm	0.343	0.359	0.348
5. Occupation*Firm interactions	0.298	0.320	0.312
6. Total between job cells	0.642	0.679	0.660
7. Within job cells	0.358	0.321	0.340
8. Total	1.000	1.000	1.000
C. Zhaopin Data			
1. Occupation	0.017	0.031	0.017
2. Firm	0.292	0.279	0.322
3. Joint occupation and firm	0.026	0.031	0.033
4. Total, occupation plus firm	0.334	0.341	0.373
5. Occupation*Firm interactions	0.304	0.327	0.294
6. Total between job cells	0.639	0.668	0.667
7. Within job cells	0.361	0.332	0.333
8. Total	1.000	1.000	1.000
D. Computrabajo Data			
1. Occupation	0.042	0.041	0.016
2. Firm	0.120	0.110	0.162
3. Joint occupation and firm	0.017	0.016	0.022
4. Total, occupation plus firm	0.179	0.168	0.199
5. Occupation*Firm interactions	0.100	0.103	0.093
6. Total between job cells	0.279	0.271	0.292
7. Within job cells	0.721	0.729	0.708
8. Total	1.000	1.000	1.000

Appendix 3: Additional Evidence on the Negative Skill-Targeting Relationship

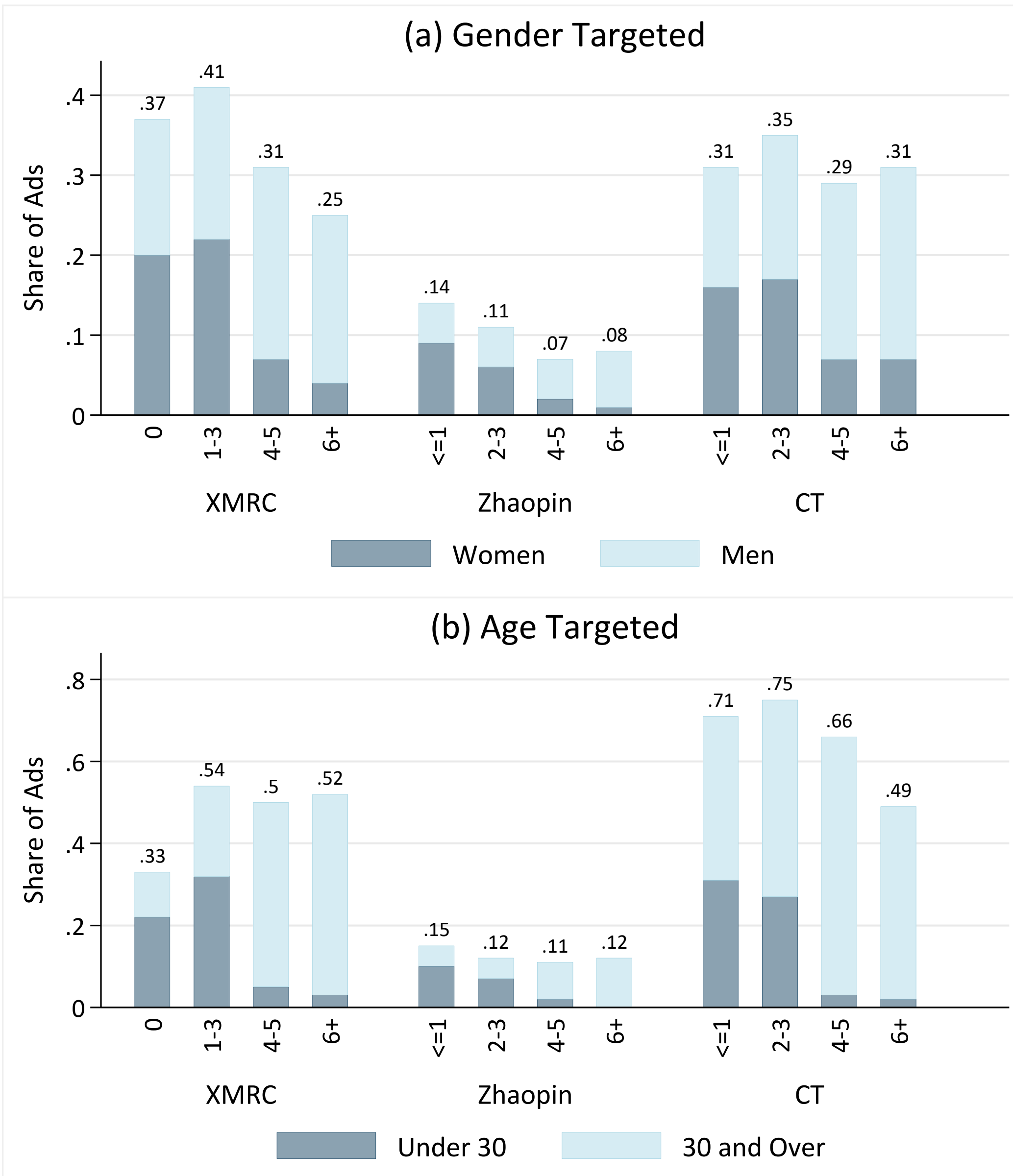
A3.1: Figures

Figure A3.1: Share of Ads that are Age and Gender Targeted, by the Job's Posted Wage



Notes: Wage bins are labeled in thousands of RMB or Mexican pesos per month. CT refers to the Computrabajo data.

Figure A3.2: Share of Ads that are Age and Gender Targeted, by the Job's Experience Requirement



Notes: Experience bins are labeled in years. CT refers to the Computrabajo data.

A3.2: Regression Analysis of Age- and Gender-Targeting

To see whether the relationships in Figures 1, A3.1 and A3.2 can be accounted for by observable differences between skilled and unskilled job ads, Tables A3.1-A3.4 present identically-specified regressions for each of our datasets. In columns 1-4 the dependent variable is an indicator for whether the ad is gender-targeted (regardless of direction); in columns 5-8 the outcome is whether the ad is age targeted (regardless of what age is requested). In other words, if P^F is an indicator for whether the ad specifically requests women, and P^M is an indicator for whether it requests men, then the dependent variable in columns 1-4 is just $P^F + P^M$, which equals either zero or one. In columns 5-8, the dependent variable equals one if the ad specifies both a maximum and minimum age, and zero otherwise. All the regressions use education requirements as our measure of job skill.[2]

Four specifications are presented for each of these two dependent variables. The first (in columns 1 and 5) has no covariates; it reproduces the raw data in the figures for comparison. Columns 2 and 6 include occupation fixed effects in addition to the covariates shown.[3] The goal is to see whether the patterns in the figures are primarily a consequence of the type of work that is done: perhaps some types of work are highly gendered, and others not, and the latter just happen to be more skilled. Columns 3 and 7 add firm fixed effects to this specification: perhaps the skill-targeting relationship results mostly from a pattern where the firms that abstain from age and gender targeting (such as, for example, foreign-owned firms) disproportionately happen to employ skilled workers for reasons unrelated to skill *per se*.

Finally, columns 4 and 8 include fixed effects for "job cells", i.e. for the interaction of firms with occupations. Here, the estimates tell us whether the same firm, advertising at two different times for the same occupation (say, sales), is more likely to gender-target its ad when seeking a highly educated salesperson than a salesperson with less education. If the negative skill-targeting relationship persists even within job cells, this suggests that it is more likely driven by a factor that is directly related to skill levels, rather than by factors that vary across the different types of work men and women do within the same firm.[4]

While the magnitude of the skill-targeting relationship does tend to attenuate as occupation controls are added to these regressions, ads requesting a college education are significantly less likely to be both age- and gender-targeted than ads requesting the lowest education level in all but one case:

[2] Similar results are obtained if we use experience requirements, or the posted wage (when available), or when all three skill measures are included in the same regression. Due to collinearity, in the latter case a subset of the indicators is sometimes statistically insignificant.
[3] Depending on the dataset, industry and province/state/district effects may also be present, depending on relevance and availability. For example, province effects are irrelevant in our two samples from the city of Xiamen, and Zhaopin is the only dataset with an industry variable.
[4] Columns 2 and 4 of Tables A3.1-A3.4 correspond, respectively, to columns 1 and 3 of KS's Table VI, though the results for Zhaopin differ slightly due to differences in specification and sample.

age targeting in the XMRC data. We have also replicated Tables A3.1-A3.4 substituting experience and offered wages for education as the job skill measure, and including all three measures with very similar results. (In these cases, offered wages have significant negative effect on age- and gender-targeting in XMRC.)

Table A3.1: Effects of Jobs' Education Requirements on the Probability an Ad is Age or Gender Targeted, XMZYJS DATA

	Prob(Gender Targeted)				Prob(Age Targeted)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High//Tech School	-0.0664***	-0.0595***	-0.0142*	-0.0030	-0.0677***	-0.0603***	-0.0257***	-0.0104
	(0.0087)	(0.0085)	(0.0073)	(0.0103)	(0.0107)	(0.0102)	(0.0081)	(0.0115)
College	-0.2203***	-0.1896***	-0.1175***	-0.1015***	-0.1300***	-0.1079***	-0.0622***	-0.0657***
	(0.0213)	(0.0190)	(0.0145)	(0.0222)	(0.0223)	(0.0214)	(0.0150)	(0.0235)
Fixed effects	None	district, occupation	occ, firm	occ*firm	None	district, occupation	occ, firm	occ*firm
number of groups	0	64	8,897	27,633	0	64	8,897	27,633
Number of ads	141,188	141,188	141,188	141,188	141,188	141,188	141,188	141,188
Adjusted R^2	0.016	0.063	0.387	0.520	0.009	0.040	0.558	0.639

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors (in parentheses) are clustered by firm.

Table A3.2: Effects of Jobs' Education Requirements on the Probability an Ad is Age or Gender Targeted, XMRC Data

	Prob(Gender Targeted)				Prob(Age Targeted)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Some Post-Secondary Education	-0.1381***	-0.1149***	-0.1055***	-0.0796***	-0.0104	0.0072	0.0241***	0.0135
	(0.0064)	(0.0064)	(0.0073)	(0.0126)	(0.0076)	(0.0077)	(0.0075)	(0.0132)
College	-0.2266***	-0.1917***	-0.1570***	-0.1332***	-0.0942***	-0.0652***	0.0102	0.0029
	(0.0093)	(0.0096)	(0.0108)	(0.0187)	(0.0123)	(0.0122)	(0.0112)	(0.0185)
Fixed effects	None	occupation	occ, firm	occ*firm	None	occupation	occ, firm	occ*firm
number of groups	0	36	6516	20618	0	36	6516	20618
Number of ads	39,727	39,727	39,727	39,727	39,727	39,727	39,727	39,727
Adjusted R^2	0.0310	0.096	0.227	0.299	0.0040	0.029	0.342	0.408

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors (in parentheses) are clustered by firm.
Columns 2 and 6 control for log firm size and firm ownership type. Columns 2-4 and 6-8 control for number of positions advertised.

Table A3.3: Effects of Jobs' Education Requirements on the Probability an Ad is Age or Gender Targeted, ZHAOPIN Data

	Prob(Gender Targeted)				Prob(Age Targeted)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Some Post-Secondary education	-0.1260***	-0.0791***	-0.0723***	-0.0648***	-0.0951***	-0.0665***	-0.0417***	-0.0256***
	(0.0113)	(0.0062)	(0.0043)	(0.0059)	(0.0095)	(0.0058)	(0.0062)	(0.0071)
University	-0.1716***	-0.1100***	-0.0991***	-0.0889***	-0.1613***	-0.1099***	-0.0549***	-0.0341***
	(0.0113)	(0.0063)	(0.0046)	(0.0075)	(0.0088)	(0.0056)	(0.0073)	(0.0087)
Fixed effects	None	occ, ind, province	occ, province, firm	occ*firm, province	None	occ, ind, province	occ, province, firm	occ*firm, province
number of groups	0	116	73,706	258,685	0	116	73,706	258,685
Number of ads	1,051,038	1,051,038	1,051,038	1,051,038	1,051,038	1,051,038	1,051,038	1,051,038
Adjusted R^2	0.031	0.077	0.331	0.561	0.025	0.058	0.385	0.562

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors (in parentheses) are clustered by firm.
Columns 2 and 6 control for log firm size and firm ownership type. Columns 2-4 and 6-8 control for number of positions advertised.

Table A3.4: Effects of Jobs' Skill Requirements on the Probability an Ad is Age or Gender Targeted, COMPUTRABAJO Data

	Prob(Gender Targeted)				Prob(Age Targeted)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Some Post-Secondary education	0.0662***	0.0549***	0.0116	0.0030	-0.0239	0.0069	0.0076	0.0038
	(0.0184)	(0.0157)	(0.0135)	(0.0139)	(0.0157)	(0.0138)	(0.0107)	(0.0114)
University	-0.0996***	-0.0840***	-0.1112***	-0.1109***	-0.1140***	-0.0518***	-0.0148	-0.0154
	(0.0221)	(0.0166)	(0.0139)	(0.0147)	(0.0229)	(0.0155)	(0.0120)	(0.0121)
Fixed effects	None	occ*state	occ*state, firm	occ*firm, state	None	occ*state	occ*state, firm	occ*firm, state
number of groups	0	441	2,383	6,973	0	441	2,383	6,973
Number of ads	90,487	90,487	90,487	90,487	90,487	90,487	90,487	90,487
Adjusted R^2	0.018	0.065	0.197	0.243	0.014	0.071	0.326	0.357

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors (in parentheses) are clustered by firm.
Regressions without firm fixed effects include controls for log firm size and firm ownership.

Appendix 4: What Explains the Negative Skill-Targeting Relationship?

As we have noted, one interpretation of the robust negative skill-targeting effect in all four of our data sets is as a direct consequence of job skill requirements: Since higher skill levels 'raise the stakes' —making it more important to identify the best job candidate—simple screening models such as KS's predict that firms should search more broadly as jobs' skill demands (indexed by θ) rise. That said, other factors could explain the negative skill-targeting relationship as well. One of these that falls outside KS's baseline model, is an increase in the stigma associated with posting an age or gender targeted ad with the job's skill level. While we cannot definitively rule out this explanation with the data at our disposal, our conversations with job board officials suggested that very little stigma is associated with posting an age or gender targeted job ad in China, at any skill level.[5]

Within KS's model also identifies a number of other factors that are predicted to affect the use of demographic screens in the hiring process. To the extent that these additional factors covary in the right direction with a job's skill level, they could also explain the negative skill-targeting relationship. In the rest of this section we briefly discuss the possible role of these factors, which are application processing costs (*c*), the expected number of applicants per position (*N*), and the unexplained variance across jobs in their relative suitability for men versus women (σ_v).[6]

Turning first to application processing costs (*c*), the screening model predicts that *ex ante* screening should become more common as *c* rises (because discouraging one group from applying saves on application processing costs). Since both intuition and available evidence suggest that application processing costs rise with jobs' skill levels, it seems unlikely that they can account for the negative skill-targeting relationship that is present in all our data sources.[7] On the other hand, the model also predicts that *ex ante* screening should become less common as the number of applicants per job (*N*) shrinks. Thus, unmeasured covariation between skill levels and labor market tightness could explain the negative skill-targeting relationship if labor

[5] One possible test of the stigma hypothesis would be to see whether gender segregation in *actual hiring* declines with skill at the same rate as gender targeting in job ads. If it declines at the same rate, then firms are probably not substituting internal gender and age based filtering of applications for profiling as skill levels rise. Unfortunately, we do not have data on which employees are hired in any of our four datasets.

[6] A fourth possibility suggested by KS's model is that the idiosyncratic variation in applicant quality (σ_ε) is higher at higher skill levels. Empirically, this is very hard to distinguish from the direct effects of skill demands (θ). (The distinction hinges on whether there is a bigger difference between the quality of a good and bad lawyer than a good and bad legal assistant, *over and above* the difference that follows from the jobs' skill demands.) Interested readers should consult KS.

[7] See Table I in Barron and Bishop (1985). In their employer survey, the total person-hours spent by company personnel recruiting, screening, and interviewing applicants to hire one individual ranged from 7.08 for blue-collar workers to 16.99 for managerial personnel.

markets for skilled workers are on average tighter. Again, the intuition is simple: why would you rule out an entire group of applicants when applicants on the whole are scarce?

A final possibility suggested by the KS model is that σ_v, the variance *across jobs* in women's (or older workers') comparative advantage, might be greater in less-skilled than in more-skilled jobs. In other words, perhaps age and gender 'matter less' for performance in skilled than in unskilled jobs. For example, if men and women are more different physically than mentally, and if age affects the performance of physical tasks more than mental tasks, jobs requiring manual labor might be more age- and gender-specialized than other jobs.[8] A related hypothesis applies this same logic to employers, co-workers' or customers' *tastes*: perhaps skilled co-workers, as well as employers or customers of skilled workers, 'care less' about the age and gender of people they interact with. In their paper, KS devise a simple test between the variance (σ_v)-based and skill (θ)-based explanations of the negative skill-targeting relationship, based on the fact that the models have different predictions for the effects of skill in jobs that are highly gendered (i.e. jobs where more than half the ads explicitly requests one of the two genders) than in other jobs.

To see the logic of this test, consider the cross-job-ad variance in the unobserved component of employer' preferences towards men, $P^M - P^F$. If this variance (σ_v^2 in KS's model) declines with job skill requirements, that could account for the negative skill-targeting relationship in all our data sources: at high skill levels, gender doesn't 'matter as much' for productivity, or isn't something that employers, co-workers or customers care so much about. If declining σ_v with skill explains the negative skill-targeting relationship, KS (2013) showed that skill increases in highly gendered jobs, i.e. in jobs where more than half the ads explicitly favor one gender at low skill levels, the share of ads directed at the dominant gender should rise even further as we increase skill requirements in those jobs. If, however, the direct effect of skill requirements (θ) is the main explanation, the share of jobs that are gendered in either direction should fall with skill in these jobs, as they do in other jobs.

To assess these differential predictions, we identified job types in which more than half the ads requested a particular gender at low skill levels (high school or less), and examined the levels of gender targeting in these jobs at higher skill levels. The results, shown in Table A4.1, are much more consistent with a direct effect of skill levels (θ): for the most part, gender-targeting declines with skill requirements, even in these highly-gendered jobs.

---

[8] Note that effects of this nature would need to occur within firm*occupation cells to explain our results. KS also show that the skill-targeting effect persists when all jobs requiring manual labor are dropped from the sample.

Table A4.1: Effects of Education Requirements on Advertised Gender Preferences in Highly Male and Highly Female Jobs

Education requirement	Share of ads requesting men in "highly male" jobs	Share of ads requesting women in "highly female" jobs
(a) XMZYJS Data		
High school or less	0.794	0.703
Some postsecondary	0.724***	0.726
University	0.588**	0.610***
(b) XMRC Data		
High school or less	0.505	0.566
Some postsecondary	0.234***	0.533**
University	0.364	0.360***
(c) Zhaopin Data		
High school or less	0.640	0.815
Some postsecondary	0.534***	0.637***
University	0.509**	0.517***
(d) Computrabajo Data		
High school or less	0.545	0.600
Some college	0.601	0.700
University	0.355***	0.750

Notes: ***, **, * refer to differences from first row, significant at 1%, 5% and 10%, respectively.

"**Highly Male**" Jobs are:
Security/fire-fighting occupation in XMZYJS;
Cars, ships and trucks occupation in XMRC;
Technical workers, maximum age 25 or over in Zhaopin;
Technical occupation, age 35 or higher in Computrabajo.

"**Highly Female**" Jobs are:
The administration occupation in XMZYJS;
The administration occupation in XMRC;
Jobs that request beauty, maximum age under 25 in Zhaopin;
The medicine/health care occupation, age below 25 in Computrabajo.

Appendix 5: Additional Evidence on the Age Twist

A5.1: Top 100 Job Title Words (XMRC)

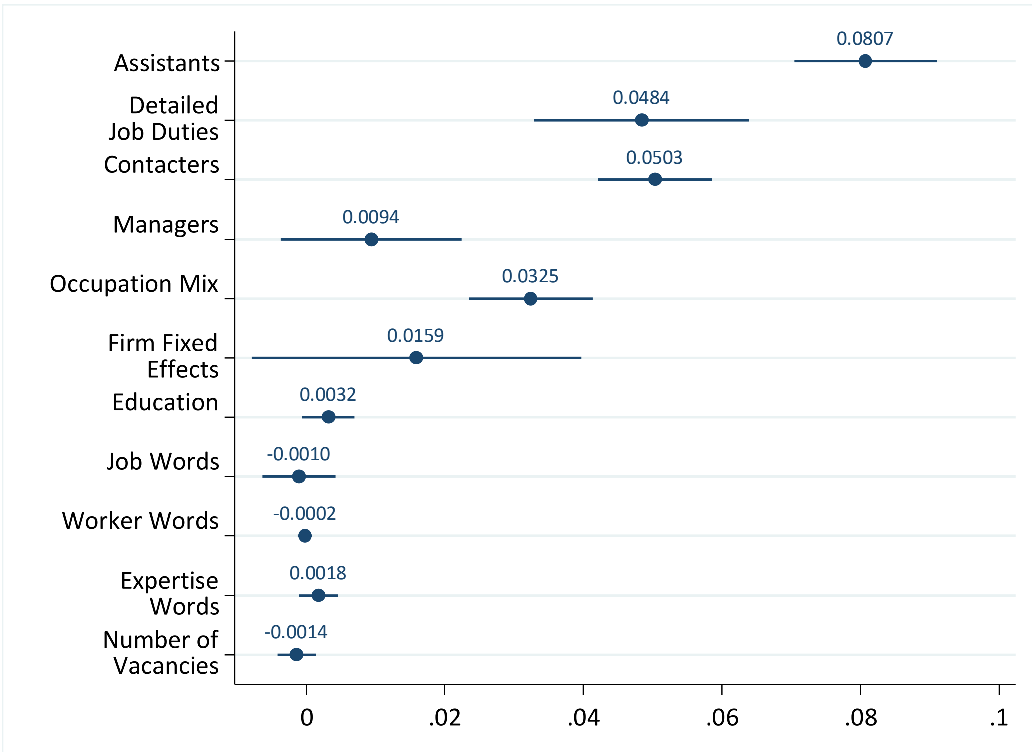
Table A5.1: 100 Most Frequent Job Title Words in XMRC
(after Encoding Job Ranks, and Indicators for
Helping, Expertise, Customer Contact Words)

administration, engineer, HR, designer, business, warehousing, accounting, division, QC, business_person, driver, general, manufacturing, marketing, IT, international_trade, finance, procurement, worker, planning, operating, molding, electronic, security, service, network, internet, shop, machine, maintenance, statistician, web, office, merchandising, injection, production, development, material, document, project, training, construction, advertising, product, promotion, delivery, repair, recruiting, truck, art, fitter, copy, certificate, software, structure, cleaning, electrical, plumber, apprentice, cost, logistics, invitation, mechanical, investment_attraction, chef, RandD, workshop, general_affairs, custom, goods, equipment, regional, teacher, programmer, implementation, cutting, inspection, welder, telephone, packaging, garment, restaurant, milling, on_site, post_sale, system, declaration, editor, brand, car, auditing, budget, legal, e_commerce, grinding, room, lathe, control, translator, Xiamen.

Note: the most frequent words are listed first.

A5.2: Age Twist Regressions: Robustness Checks

Figure A5.1: 'Explained' Components of the Age Twist, top 50 words, XMRC



Notes: Figure A5.1 replicates Figure 3 using only the 50 most common job duty words.

Table A5.2: Effects of Desired Age on Firms' Gender Preferences ($P^M - P^F$), Substituting 880 Mutually Exclusive Job Duty Categories for Duty-Word Fixed Effects, XMRC Data

	(1)	(2)	(3)	(4)	(5)	(6)
Age30plus	0.4420***	0.3723***	0.2758***	0.2727***	0.1883***	0.1727***
	(0.0148)	(0.0138)	(0.0143)	(0.0143)	(0.0142)	(0.0238)
Some Post-Secondary		-0.0928***	-0.1014***	-0.0920***	0.0582***	0.0605**
		(0.0150)	(0.0141)	(0.0145)	(0.0154)	(0.0268)
University		0.0350	0.0341	0.0423*	0.1806***	0.1366***
		(0.0244)	(0.0239)	(0.0241)	(0.0254)	(0.0432)
Managers			-0.0339**			
			(0.0165)			
Assistants			-0.4751***			
			(0.0214)			
Contacters			-0.4146***			
			(0.0210)			
Experts			-0.0024			
			(0.0200)			
High_Mgr				-0.0012	0.1470***	0.1326**
				(0.0318)	(0.0364)	(0.0571)
Manager				-0.0453**	0.0955***	0.0636*
				(0.0184)	(0.0249)	(0.0385)
Supervisor				-0.1771***	0.0844*	0.0534
				(0.0392)	(0.0442)	(0.0660)
Assistant				-0.4300***	-0.3019***	-0.3433***
				(0.0265)	(0.0299)	(0.0450)
Clerk				-0.5750***	-0.4590***	-0.5034***
				(0.0249)	(0.0276)	(0.0478)
Secretary				-0.5724***	-0.5037***	-0.6442***
				(0.0528)	(0.0555)	(0.0910)
Helper				0.4207	0.9116***	0.7355***
				(0.4218)	(0.0637)	(0.1160)
Front_Desk				-0.3831***	-0.3076***	-0.3820***
				(0.0285)	(0.0294)	(0.0514)
Customer_Service				-0.3848***	-0.3136***	-0.4195***
				(0.0396)	(0.0421)	(0.0686)
Teller				-0.4494***	-0.5221***	-0.5086***
				(0.0382)	(0.0470)	(0.0782)
Cashier				-0.5554***	-0.4441***	-0.3570***
				(0.0762)	(0.0794)	(0.1279)
Sales				-0.2949***	-0.0331	0.0843
				(0.1008)	(0.1696)	(0.2666)
Representative				0.0954	0.1551**	0.1042
				(0.0627)	(0.0635)	(0.0930)
Number of ads	9,884	9,884	9,884	9,884	9,884	9,884
R^2	0.0891	0.2678	0.3334	0.3441	0.5124	0.7329

*** p<0.01, ** p<0.05, * p<0.1. Standard errors (in parentheses) are clustered by firm.

Notes to Table A5.2:

Specifications without job title controls:

Column 1: no controls

Column 2: education requested, number of vacancies, firm size and location, plus 36 occupation fixed effects.

Specifications with job title controls:

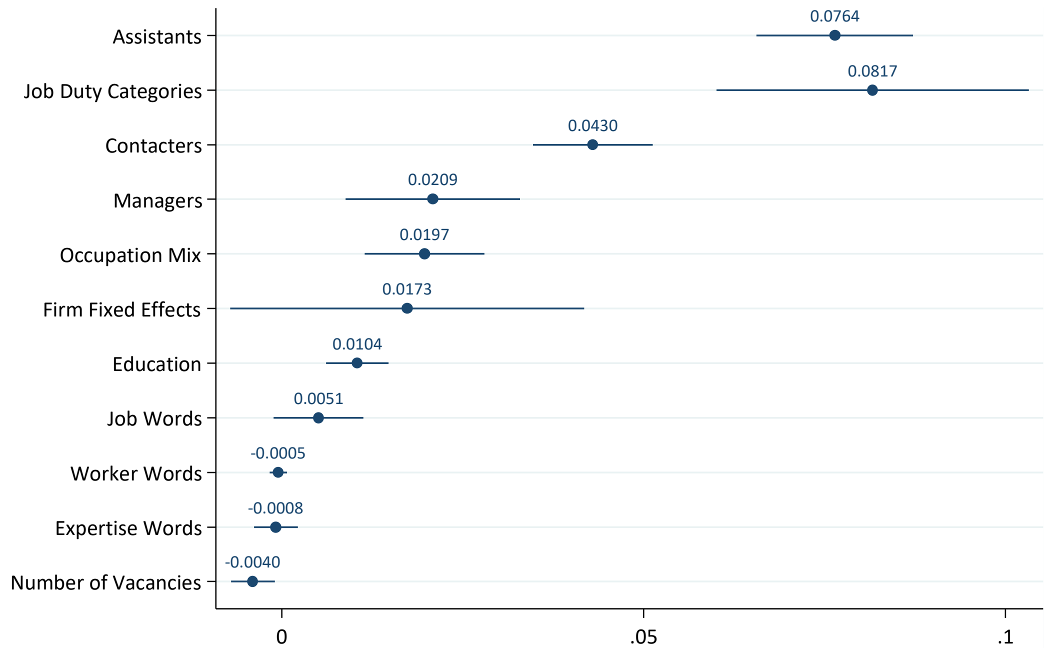
Column 3: adds four broad indicators from the job title: Managers, Assistants, Contacters and Experts.

Column 4: replaces these four broad indicators by dummies for individual title words in each category. Also adds also adds the miscellaneous worker words and job words extracted from the title.

Column 5: adds fixed effects for 880 mutually exclusive job duty categories. Each category is a unique combination of all the title words not already encoded in the Managers, Assistants, Contacters, and Experts words and the miscellaneous worker words and job words. Of these 880 categories, 690 were retained in the regression after eliminating collinear words.

Column 6: adds firm fixed effects. Of these 880 categories, 647 were retained in the regression after eliminating collinear words.

Figure A5.2: 'Explained' Components of the Age Twist, 880 Job Duty Categories, XMRC Data



Notes: Figure A5.2 replicates Table 3, controlling for 880 mutually exclusive job duty categories.

Table A5.3: Effects of Desired Age on Firms' Gender Preferences ($P^M - P^F$), Adding Controls for Beauty Requests—Zhaopin and Computrabajo datasets

	(1)	(2)	(3)	(4)
	ZHAOPIN Data		**COMPUTRABAJO Data**	
Age30plus	0.1357***	0.1150***	0.1254***	0.1160***
	(0.0168)	(0.0176)	(0.0135)	(0.0136)
Beauty requested?		-0.1951***		-0.1741***
		(0.0232)		(0.0277)
Photo requested?				-0.1587***
				(0.0261)
Fixed effects	occ*firm, province	occ*firm, province	occ*firm, state	occ*firm, state
Number of groups	50,048	50,048	5,774	5,774
Number of ads	134,100	134,100	65,516	65,516
Adjusted R^2	0.8114	0.8144	0.2984	0.3062

Notes: Columns (1) and (3) repeat column 4 in Tables 5 and 6 respectively. Columns (2) and (4) add a beauty control to those specifications.

In Computrabajo, the beauty request indicator equals one if the ad either requested beauty or requested a photo.

Share of the twist explained by the beauty/photo controls equals $\left(\frac{.1357-.1150}{.1357}=\right)15.25$ percent in Zhaopin and $\left(\frac{.1254-.1160}{.1254}=\right)7.50$ percent in Computrabajo.

A5.3: Size of the Age Twist by Occupation

Figures A5.3-A5.6 show the size of the age twist by occupation in each of our four datasets. In all these figures, the age twist is defined as the difference in employers 'preferences towards men' $(P^M - P^F)$ between ads requesting workers under versus over 30 years of age. [9]

Figures A5.3-A5.6 show that the age twist is positive in almost all occupations in all datasets, but varies substantially in magnitude across occupations. Three occupations stand out as having consistently high twists: administration (which includes receptionists, executive assistants and other office assistants), customer service, and management.[10]

---

[9] To keep the figures manageable and more comparable across datasets, and to increase the precision of the within-occupation twist measures, this analysis uses occupation groups that are more aggregated than the ones used in our regression analyses. Specifically, compared to the regressions, the number of categories is reduced from 58 to 19, 36 to 23, 39 to 26 and 14 to 13 in XMZYJS, XMRC, Zhaopin and Computrabajo respectively.

[10] The one exception is that management has a low twist in XMRC.

Figure A5.3: Age Twist by Occupation: XMZYJS Data

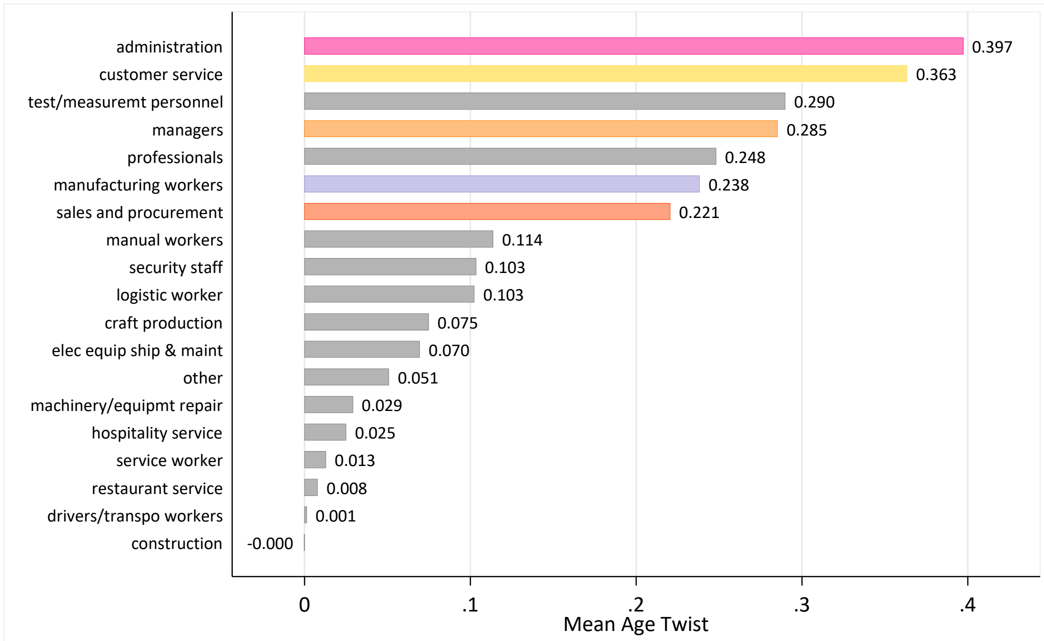


Figure A5.4: Age Twist by Occupation: XMRC Data

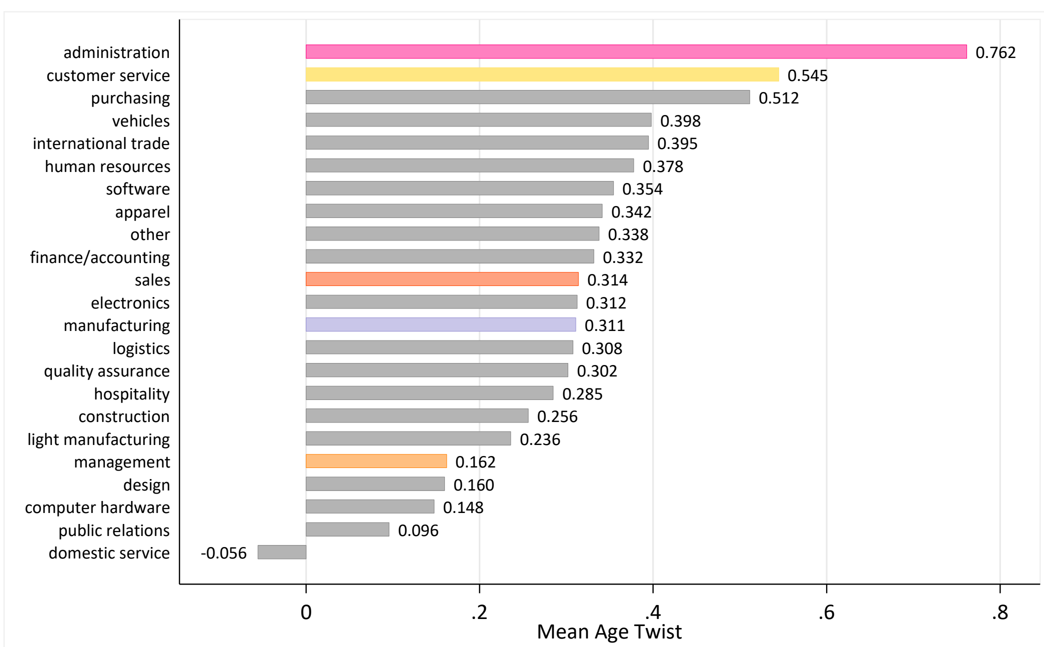


Figure A5.5: Age Twist by Occupation—Zhaopin Data

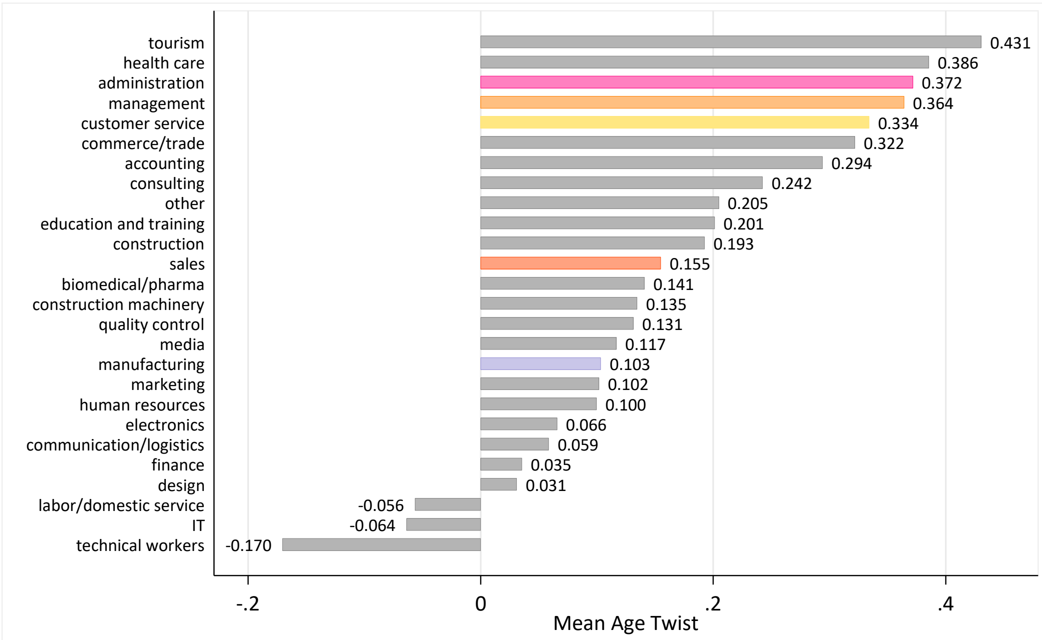
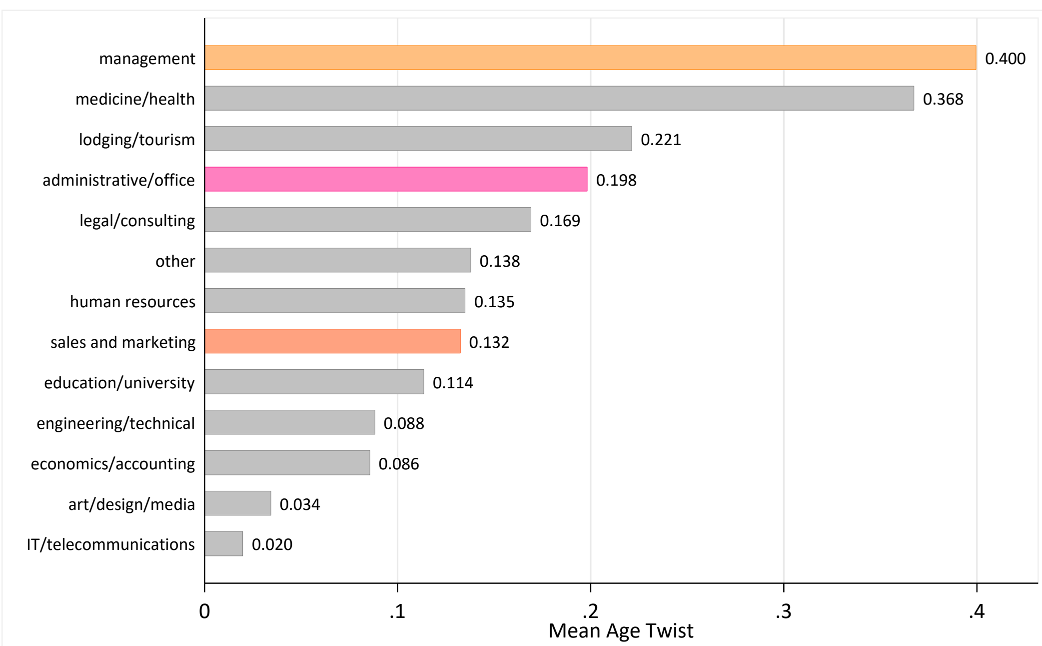


Figure A5.6: Age Twist by Occupation—Computrabajo Data



Notes to Figures A5.3-A5.6:

The age twist is the difference in employers 'preferences towards men' ($P^M - P^F$) between ads requesting workers under versus over 30 years of age. Asterisks (*) indicate that the occupation's title suggests a customer contact component.

In addition to administration, customer service, and management, Figures A5.3-A5.6 also flag two large occupations —sales and manufacturing workers—both of which have about average-sized twists. In part due to their large size, almost half of the age twist in all four datasets occurs within these five occupations.

A5.4: Occupation-Level Regressions for the Size of the Age Twist

TABLE A5.4: Effects of Mean Beauty Requests on the Magnitude of an Occupation's Age Twist, Zhaopin and Computrabajo Data

	(1) **Zhaopin Data**	(2) **Computrabajo Data**
Mean beauty demand	0.8061***	1.2154*
	(0.2034)	(0.6511)
Number of Occupations	26	13
Adjusted R^2	0.370	0.172

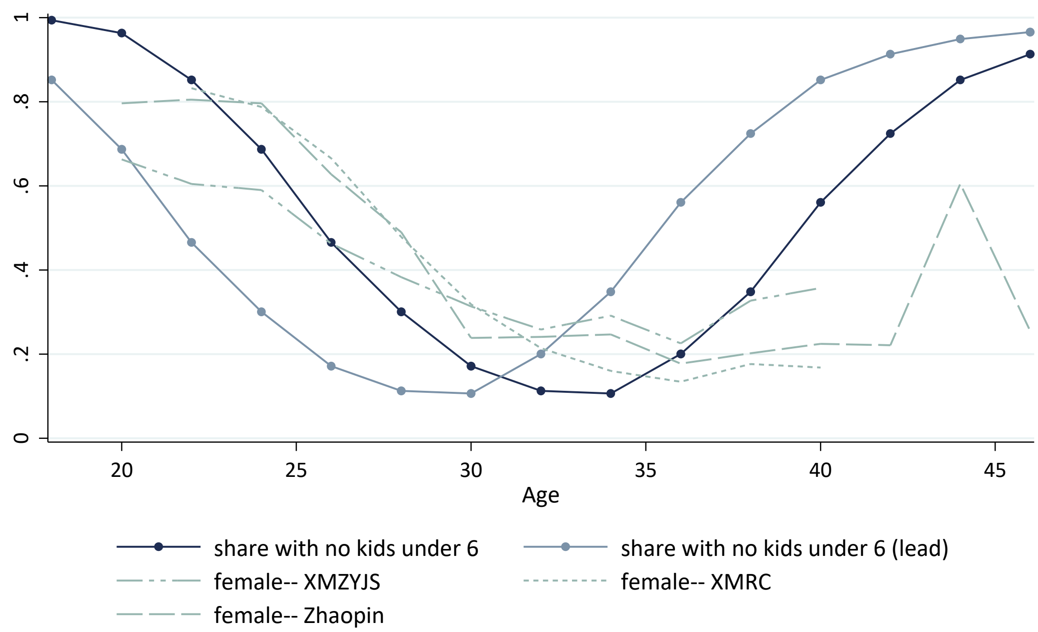
Notes: *** p<0.01, ** p<0.05, * p<0.1. The age twist is the difference in preference towards men $(P^M - P^F)$ between ads requesting workers 30 and older and ads requesting younger workers in the occupation. Regressions are weighted by the number of ads in the occupation. Computrabajo beauty indicator=1 if the ad requests beauty or a photo.

Table A5.4 estimates simple cross-occupation regressions in the Zhaopin and Computrabajo datasets. The dependent variable in these regressions is the size of the occupation's age twist, from Figures A5.3 and A5.4. The single regressor is a continuous variable equal to the share of ads in the occupation that explicitly request an attractive applicant.[11] In both datasets, occupations with a higher demand for beauty exhibit larger *shifts* in employers' preferences *away from women as workers age* (p<.01 in Zhaopin and p<.10 in Computrabajo). Quantitatively, the R^2s in Table A5.4 indicate that gendered employer preferences for beauty can account for 37 percent of the cross-occupational variation in the age twist in the Zhaopin data, and for 17 percent in the Computrabajo data.

[11] To economize on degrees of freedom, Table A5.4 uses an aggregate indicator of beauty demand— whether the ad requested beauty or a photo. If these are entered separately, both have positive coefficients but the photo indicator has the stronger and more significant effect.

A5.5: Additional Evidence on Motherhood and the Timing of the Age Twist

Figure A5.7: Female Share of Gendered Job Ads versus Share of Women who have (or will have) *a child under age 6*, China



Notes: Dashed lines show the share of gendered job ads that request women in each job board, reweighted to match the education distribution of urban Chinese women. Solid lines show the share of urban Chinese women who have a child under age six at that age, or within the following two years. Marital status and education information is for working age urban women in the 2005 Census 1% microdata sample.

Appendix 6: Labor Market Activity over the Life Cycle in Urban China and Mexico

Figure A6.1: Share of the Employed Population in the Private Sector, China

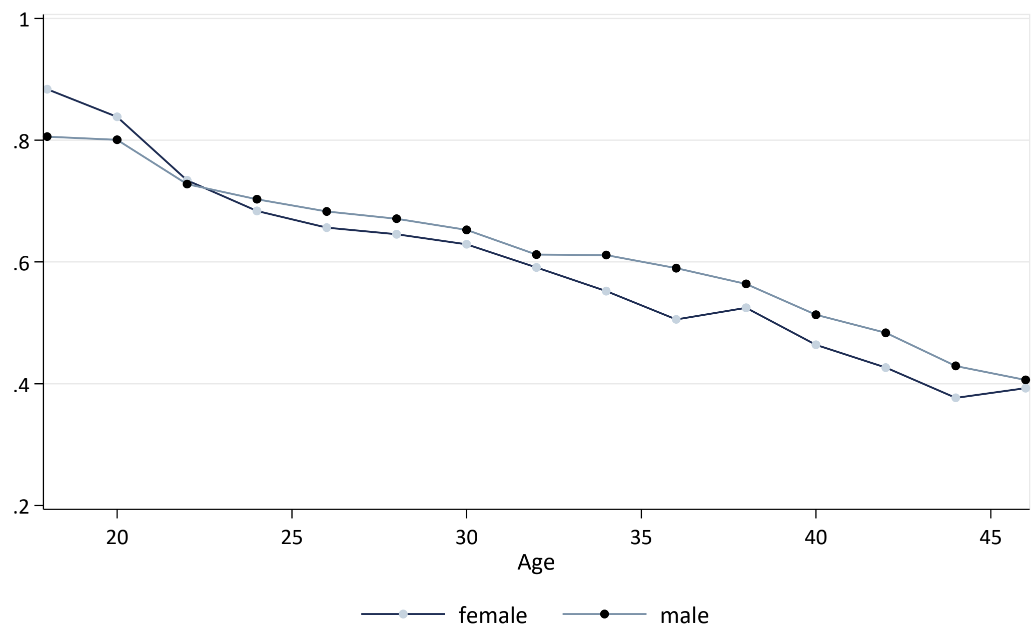


Figure A6.2: Share of the Private-Sector Employed Population with Contract of Two Years or Less, or No Contract, China

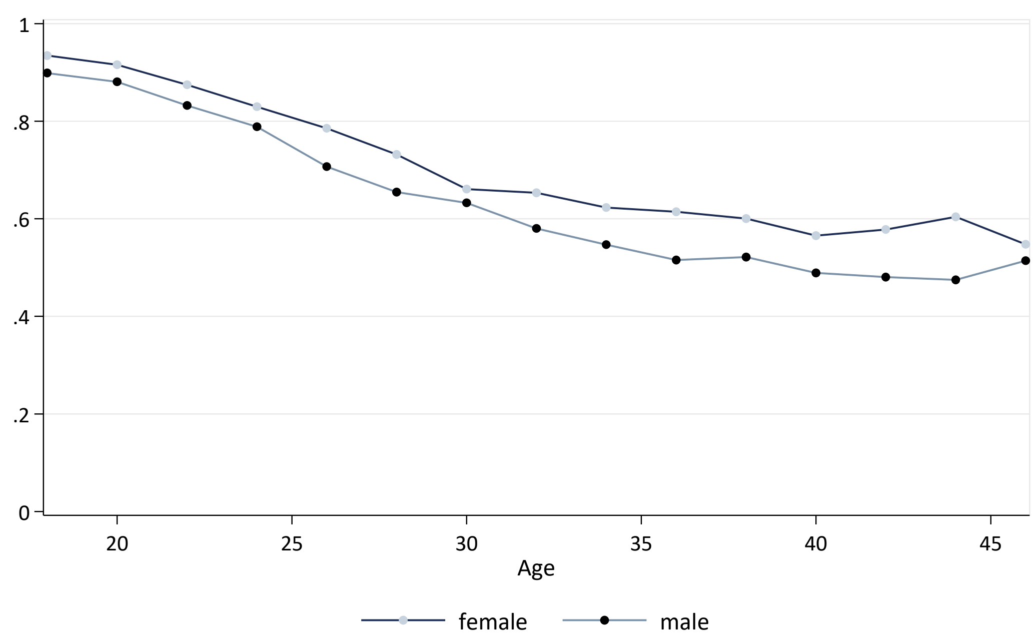


Figure A6.3: Mean Weekly Hours of the Private-Sector Employed Population, China

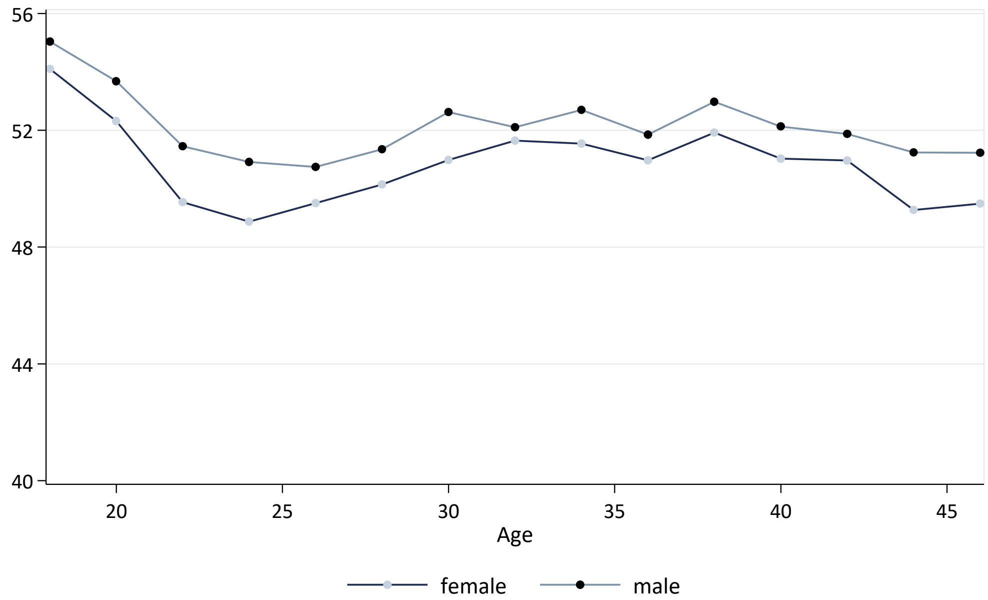


Figure A6.4: Part-Time Work-- Share of the Private-Sector Employed Population Working Less Than 30 Hours per Week, China

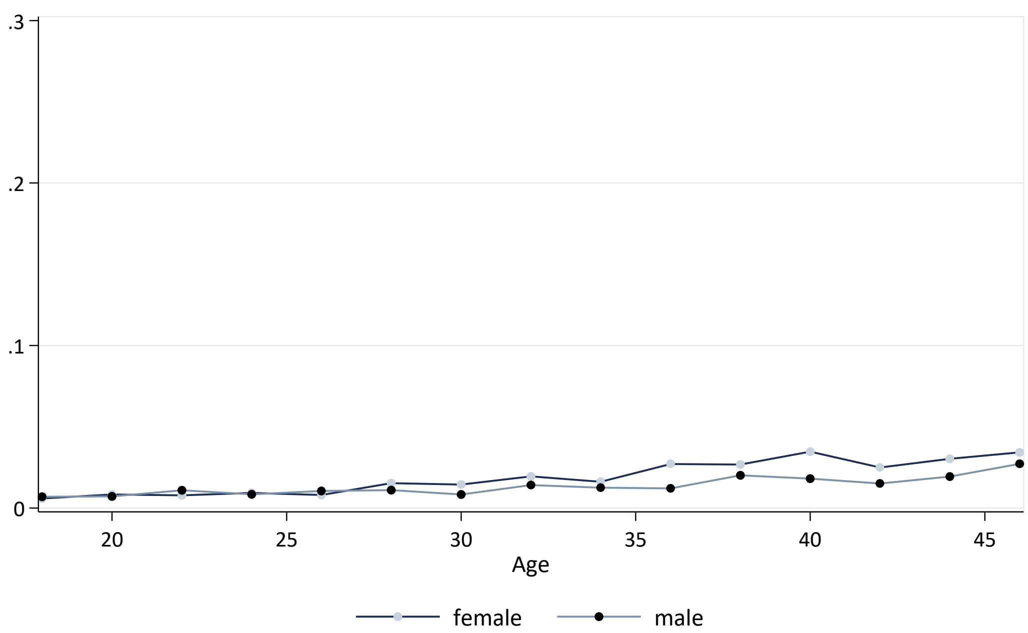


Figure A6.5: Long Work Hours-- Share of the Private-Sector Employed Population Working 50 or More Hours per Week, China

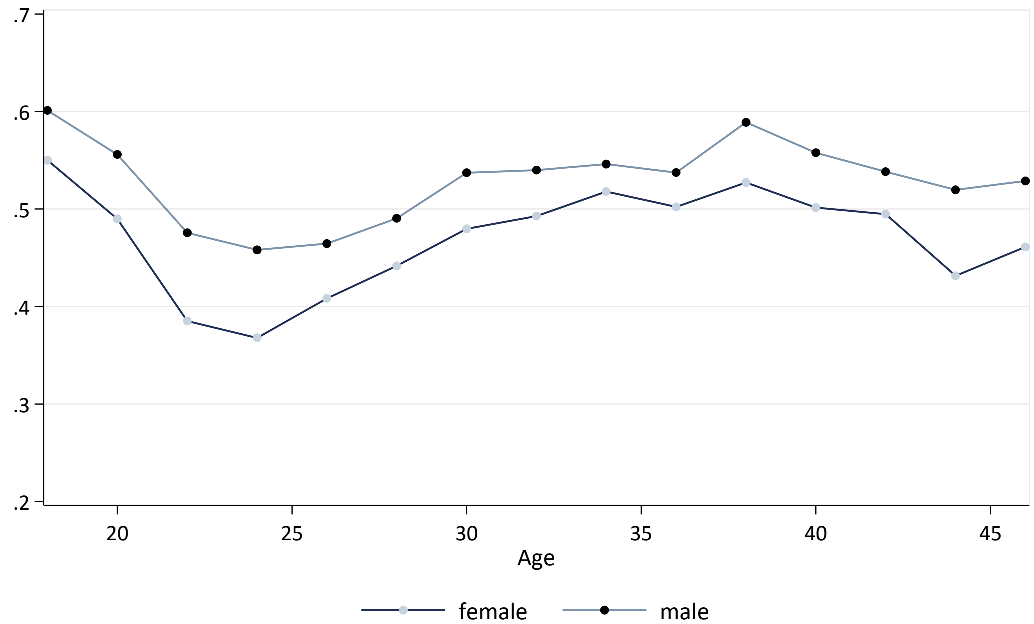
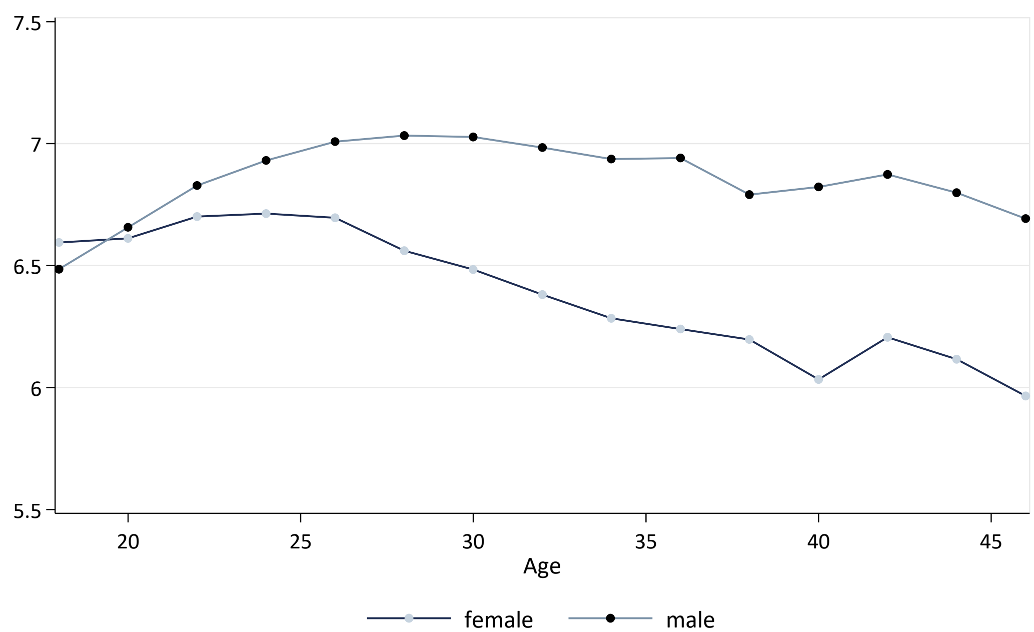


Figure A6.6: Average Log Monthly Earnings of Private-Sector Full-Time Workers, China



Notes: High rates of educational upgrading substantially flatten age-wage profiles in China relative to Mexico and other countries.

Figure A6.7: Share of the Employed Population in the Private Sector, Mexico

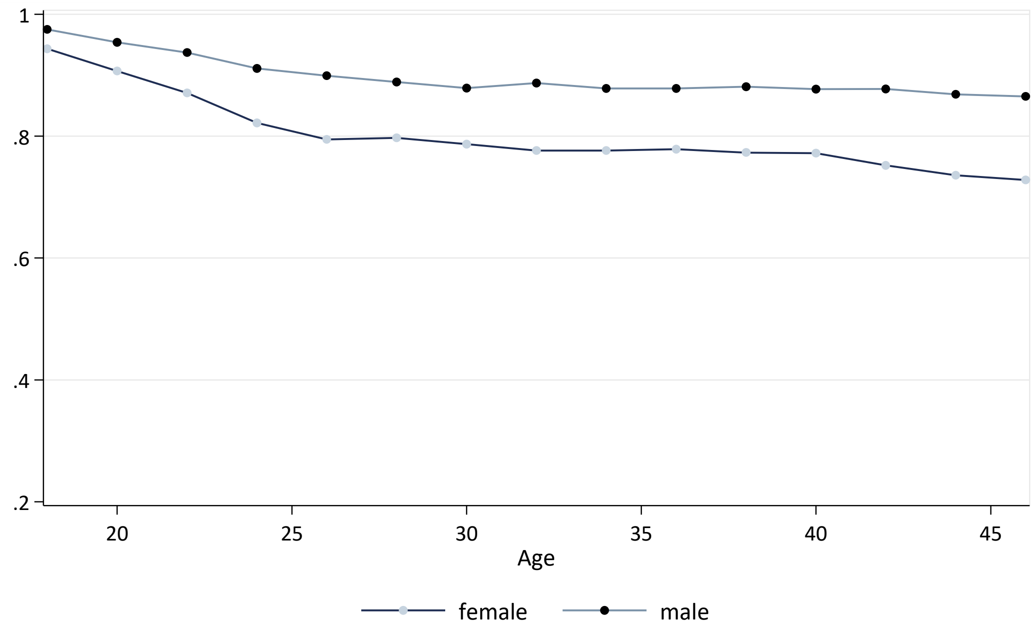


Figure A6.8: Share of Private-Sector Employed Population in Urban Areas who have Job Tenure of Two Years or Less, by Age and Gender, Mexico

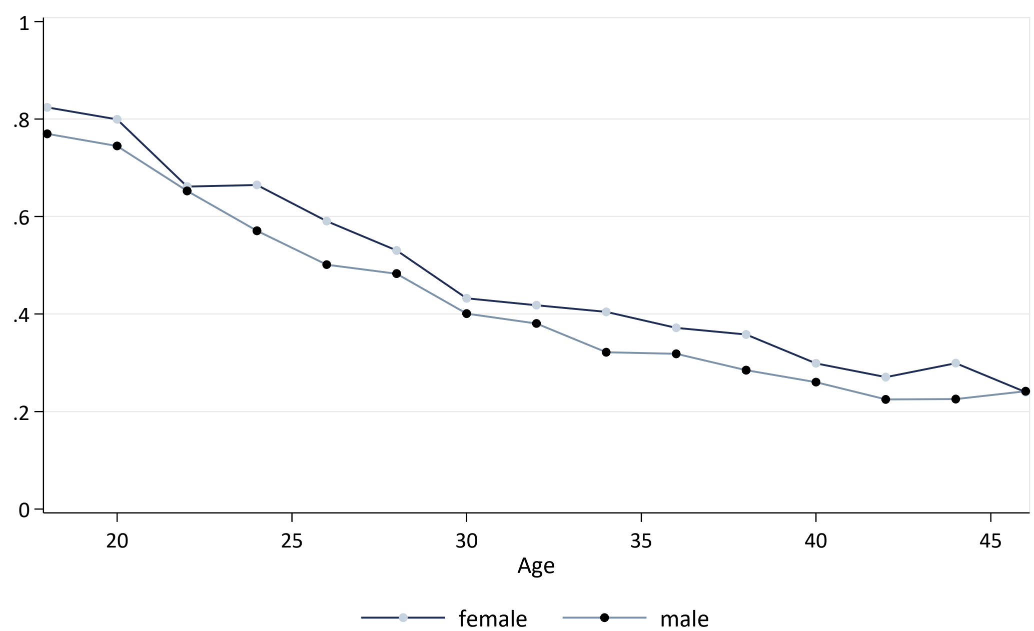


Figure A6.9: Mean Weekly Hours of the Private Sector Employed Population, Mexico

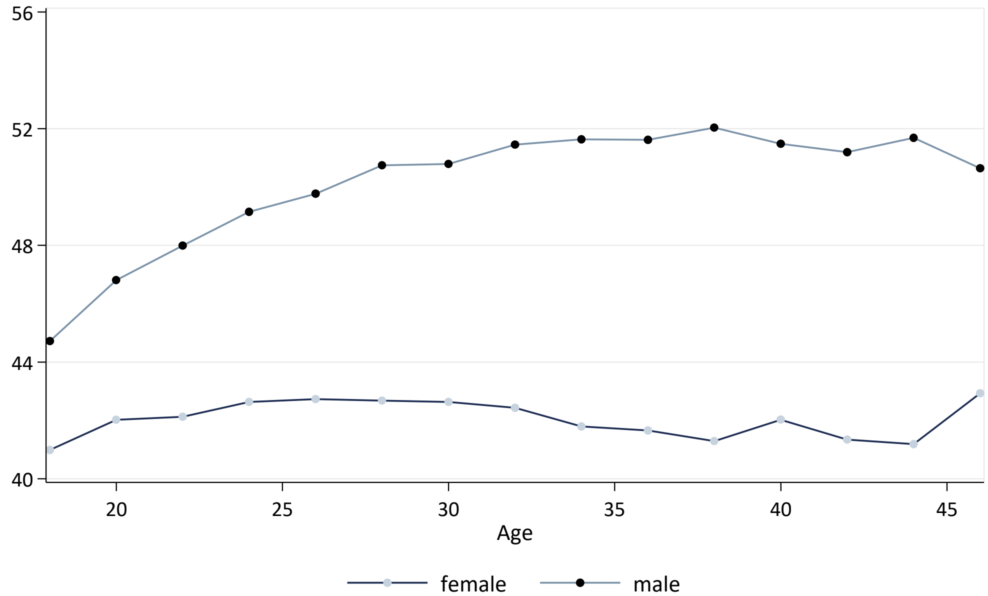


Figure A6.10: Part-Time Work-- Share of The Private-Sector Employed Population Working Less Than 30 Hours Per Week, Mexico

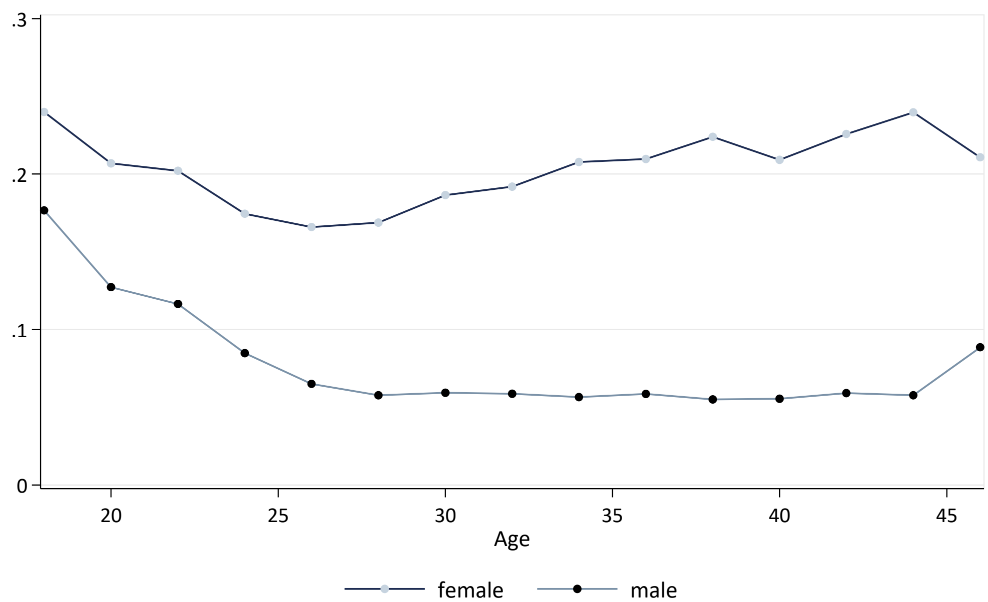


Figure A6.11: Long Work Hours-- Share of The Private-Sector Employed Population Working 50 or More Hours per Week, Mexico

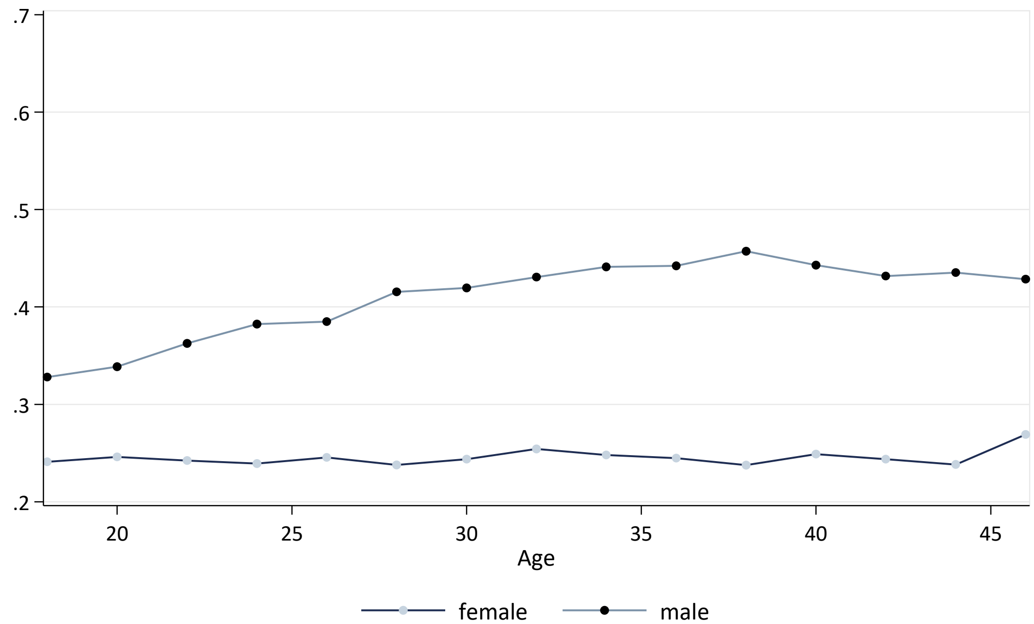
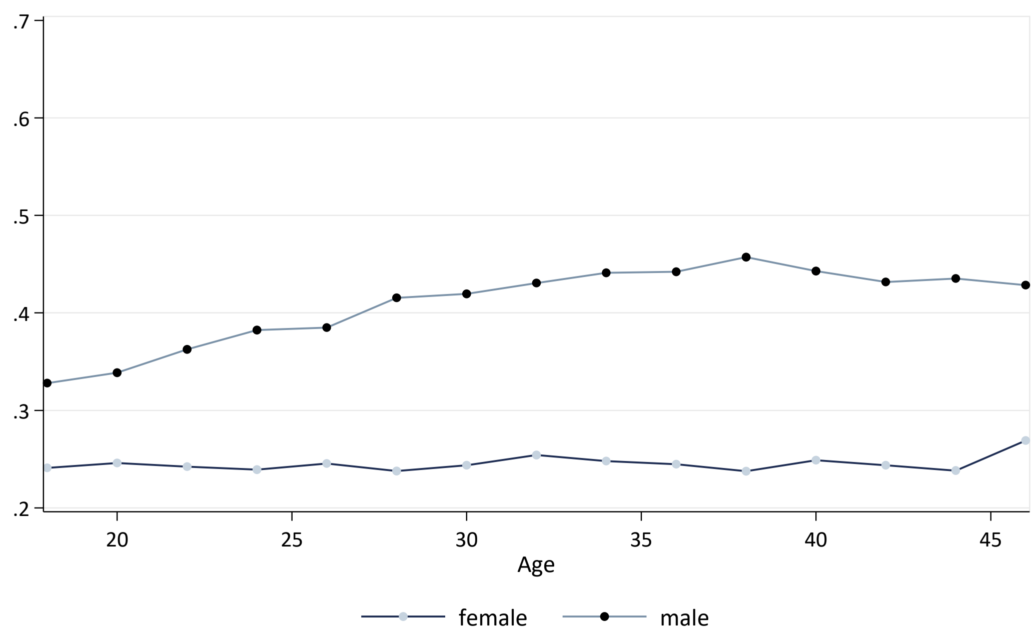


Figure A6.12: Average Log Monthly Earnings of Private-Sector Full-Time Workers, Mexico



Sources and Definitions for Figures A6.1-6.12:

All statistics for China are authors' calculations from the Chinese 2005 Census 1% sample. Sample is restricted to individuals aged 18-45 living in urban or suburban China (i.e. excluding rural China), with valid age, gender, employment and employer type information.

Statistics for Figure A6.8 are authors' calculations from 2010 Q1 National Survey of Occupations and Employment (ENOE). Sample restricted to urban population, age 16-45. All other statistics for Mexico are data from the 2010 Census microdata file, and refer to urban persons aged 16-45 who reported working at least one hour the week before the census interview. Private-sector employed are those who reported working in the private sector (i.e. excluding legislative bodies, federal, state, and municipal administration, international organizations, public schools, public hospitals, central banking, and other public sector services). Urban population is defined as areas ("localidades") with more than 100,000 inhabitants.