To Veil or Not to Veil? Assessing the Removal of Headscarf Ban in a Muslim Country*

Wenqi Lu† Ekin Yurdakul‡

October 25, 2025

Abstract

This paper examines how removing identity-based institutional restrictions affects women's economic participation by exploiting the 2013 repeal of a longstanding headscarf ban in Turkey. We combine two nationally representative surveys and use statistical matching and machine-learning models to predict women's veiling status, to identify treatment effects by comparing veiled and non-veiled women's labor market outcomes in a difference-in-differences framework. The repeal led to a significant rise in public sector employment among veiled women, driven by both higher employment rates and shifts away from self-employment and unpaid family work. In contrast, non-veiled women experienced a decline in public sector jobs, suggesting a substitution effect, with no evidence of reduced efficiency in the public sector after the repeal. Consistent with this pattern, effects are concentrated among more educated veiled women, the group most likely to qualify for public positions. We find no differential effects across regions with high and low veiling prevalence or ruling-party vote shares, suggesting that institutional access, rather than local acceptance or political favoritism, drives the response.

JEL: J16, J12, J22, K31

Keywords: headscarf ban, identity-based policy, women's economic participation

---

*We are particularly grateful to Luca Paolo Merlino for extensive comments and valuable guidance. We also thank Laurens Cherchye, Giuseppe Sorrenti, Paula Gobbi, Sunčica Vujić, Bahadır Dursun, Thomas van Huizen and Marion Leturcq for helpful feedback and suggestions. All remaining errors are our own.

†University of Antwerp & Université Libre de Bruxelles and Aix-Marseille Univ. (Aix-Marseille School of Economics), CNRS & EHESS. E-mail: wenqi.lu@uantwerpen.be

‡Corresponding author. Utrecht School of Economics, Utrecht University. e.yurdakul@uu.nl

1 Introduction

Economic participation is often shaped by the interaction between cultural identity and institutional rules (Akerlof and Kranton, 2000; Fernández and Fogli, 2009; Alesina et al., 2013; Alesina and Giuliano, 2015; Oh, 2023). When institutions regulate how identity can be expressed, for example, by veiling bans or dress codes, they may alter individuals' incentives and opportunities in the labor market. Recent research highlights the importance of identity-based policies in shaping behavior and welfare (Carvalho, 2013; Fouka, 2020; Saleh and Tirole, 2021; Abou Daher et al., 2025). This raises a central question: when legal barriers linked to identity are removed, do individuals adjust their economic choices, or do persistent identity-based norms continue to constrain participation? The answer matters for designing policies that aim to expand economic opportunity by changing formal institutions rather than deeply embedded cultural norms (Bursztyn et al., 2020; Jayachandran, 2021; Dhar et al., 2022).[1]

To provide evidence on whether easing identity-based restrictions can translate into greater economic inclusion, we study the repeal of a longstanding headscarf ban, which lifted a major institutional barrier to veiled women's employment in public institutions in Turkey. Until 2013, wearing headscarves in all public institutions was prohibited under the country's constitutionally secular framework. In October 2013, the government lifted this ban, allowing women to wear headscarves in public institutions, including government offices. Given that roughly 60% of Turkish women wear headscarves, this reform represents a rare and sharp nationwide elimination of a legal barrier directly linked to a widespread cultural practice, providing a distinctive setting to study how women's labor market outcomes respond to the lifting of an identity-based institutional constraint.

We use data from the 2010–2017 Turkish Household Labor Force Survey (HLFS) and employ a difference-in-differences strategy to estimate the effects of the policy change. A key empirical challenge in the literature on headscarf bans is the lack of direct information on veiling status in large-scale surveys, which has often limited identification strategies. To address this, we leverage the 2013 Turkish Demographic and Health Survey (DHS), which uniquely contains individual-level information on veiling together with detailed background characteristics. We integrate the two datasets using statistical matching techniques (Rubin, 1986) to estimate veiling probabilities in the HLFS based on observable characteristics. Unlike previous work that relies on broad proxies such as ethnicity or immigrant status to identify the effects of institutional restrictions regarding veiling, our approach uses individual-level survey data to impute veiling status as a proxy for treatment assignment, thereby improving the precision of group identification.[2] This enables a quasi-experimental comparison of labor market outcomes between veiled and non-veiled women before and after the policy change. Robustness checks using alternative prediction models and machine-learning algorithms (Mullainathan and Spiess, 2017; Athey and Imbens, 2019; Heller et al., 2024) confirm that the results are not sensitive to different model specifications.

Our findings reveal that following the removal of the headscarf ban, women predicted to be veiled experienced a 1.2-percentage-point increase in employment (about 4 percent relative to the sample mean) compared with their non-veiled counterparts. This rise is

[1] In this paper, we use the term cultural identity to refer to forms of identity grounded in religious or traditional values, such as veiling in Muslim societies. Cultural norms denote the informal behavioral expectations associated with that identity, for example, beliefs about appropriate roles for veiled women in public life and labor market.

[2] We first estimate each woman's probability of veiling using the 2013 DHS based on observable characteristics, and then apply nearest-neighbor statistical matching (Rosenbaum and Rubin, 1983) to impute veiling status in the HLFS. This two-step approach links women across surveys with comparable background characteristics and enables consistent treatment classification (see Ridder and Moffitt (2007) for related methods).

driven primarily by a substantial increase in public sector employment among veiled women, whose participation in such jobs was nearly zero before the reform. At the same time, we observe notable declines in self-employment and unpaid family work, suggesting that the repeal reshaped employment choices among women already in the workforce. Consistent with this shift, the probability of informal employment fell, while the likelihood of holding a permanent contract increased. Following the repeal, veiled women's average weekly working hours rose by about half an hour per week, reflecting higher employment rates rather than longer hours among those already working. Examining the occupational distribution within the public sector, we find the largest gains in professional and associate professional roles, such as teachers, nurses, and administrative staff, which typically require some level of educational attainment.

Interestingly, we also find that non-veiled women experienced a shift away from public sector jobs toward private sector employment, despite overall public sector hiring remaining stable. This pattern indicates a substitution effect: the repeal reallocated public employment opportunities from non-veiled to veiled women rather than generating new jobs. A natural question is whether this substitution reflects politically motivated hiring that may have reduced efficiency in public employment. Regional patterns suggest that the post-repeal rise in veiled women's public employment is not concentrated in politically aligned areas. Our analysis of education–occupation matching within the public sector reveals no evidence of systematic favoritism or major efficiency losses. The decline in overqualification among veiled entrants suggests that average qualification levels may have slightly decreased, but overall job–skill alignment improved or remained stable, indicating that the reform broadened access without substantially compromising workforce quality.

Two potential identification concerns deserve emphasis. The first is that predicted veiling status may be correlated with unobserved religiosity, which could independently influence women's labor supply decisions (Carvalho, 2013). Because religiosity is not observed in our main dataset, we proxy for it using pre-reform regional veiling prevalence, and find no evidence that the estimated effects are concentrated in more religious areas. Consistent with the institutional scope of the reform, the gains of the repeal are instead concentrated among relatively more educated veiled women, those more likely to access public sector jobs, while effects for less educated women are negligible. The second concern is that our identification strategy implicitly treats veiling as a stable trait. Although individual panel data on veiling are not available, we document very high regional stability in veiling rates across the 2008 and 2013 DHS waves. More importantly, we exploit two major institutional reforms, the 1997 extension of compulsory schooling and the 2002 Civil Code reform, raising the minimum marriage age to 18, and show that neither reform had any sizable effect on the probability of veiling. This evidence suggests that veiling is highly persistent and unlikely to respond to short-run policy changes. To the extent that some women may have altered their veiling status in response to the policy, it would attenuate the estimates, implying that our results should be interpreted as conservative lower bounds.[3]

Given that veiling is often associated with lower employment in both Muslim-majority and Western contexts (Ghumman and Jackson, 2010; Abdelhadi, 2019; Fernández-Reino et al., 2023), our findings shed light on whether veiled women's lower labor market participation primarily reflects individual religious identity constraints or institutional barriers (Aksoy and Gambetta, 2016; Joslin and Nordvik, 2021; Shofia, 2022; Jacquet and Montpetit, 2023). Exploiting the repeal of a long-standing headscarf ban in a secular, yet Muslim-majority, setting, we show that removing institutional restrictions on religious expression

[3]We discuss exogeneity of veiling status in detail in Section 4.

substantially increased veiled women's employment, most visibly in the public sector, where the ban had applied directly. This finding highlights that institutional constraints can play a decisive role in shaping women's economic choices. Conceptually, our evidence suggests that reducing the cost of identity expression need not hinder economic inclusion; on the contrary, relaxing institutional barriers can facilitate the integration of women previously excluded from employment due to their religious identity.

Economic theories of veiling predict that when restrictions impose significant costs on religious expression, some women may withdraw from employment rather than compromise their religious identity, while others may adjust their practices if economic incentives outweigh these costs (Carvalho, 2013). Despite the increasing prevalence of veiling regulations worldwide, systematic evidence on their economic and social consequences remains limited, largely due to the absence of direct information on veiling status.[4] A few exceptions provide mixed evidence, reflecting this theoretical ambiguity. Abdelgadir and Fouka (2020) study the 2004 French headscarf ban in schools using country of birth to proxy for Muslim identity and find that exposure to the ban significantly reduced secondary educational attainment among female students of North African origin, with long-term negative effects on labor market outcomes and family composition. Using a similar empirical approach but focusing on a different cohort, Maurin and Navarrete H (2023) exploit nationality at birth to identify Muslim students and find that an earlier 1994 ministerial circular prohibiting veiling improved educational outcomes for female students with a Muslim background.

Unlike these studies, which analyze the effects of restrictions on wearing a headscarf, our paper provides one of the first causal estimates of the consequences of lifting such a restriction in public spaces. While the proxies used by Abdelgadir and Fouka (2020) and Maurin and Navarrete H (2023)—country and nationality at birth—are plausibly exogenous and appropriate given data limitations, they remain imperfect indicators of veiling, as only around 30% of Muslim women are veiled in France (Drouhot et al., 2023). Moreover, these proxies are not informative in Muslim-majority contexts, where national or ethnic origin does not distinguish between veiled and non-veiled women. Our approach uses prediction-based statistical matching methods in a difference-in-differences framework to identify the treatment effects of lifting veiling restrictions, offering an alternative, though still indirect, strategy that is better suited to settings where the majority is Muslim.

Given the key differences in treated populations, we further examine heterogeneity by regional veiling prevalence to provide intuition across Muslim-minority and Muslim-majority contexts. Specifically, we contrast regions where veiled women constitute a minority (below 50 percent) with those where they are the majority. We find no significant differences in labor market responses between these groups, suggesting that peer effects and local social acceptance are not the main drivers of our results. This heterogeneity analysis is, of course, specific to variation in veiling prevalence within a Muslim-majority country, and any external parallels should be viewed as suggestive rather than directly comparable. Nonetheless, the pattern we document offers indicative implications for contexts such as France, where veiled women form a minority and may face discrimination. The finding that veiled women in Turkey respond similarly regardless of local veiling prevalence suggests that easing institutional restrictions can enhance participation even in environments where veiling is less

[4] Throughout the text, we use the term "veiling regulations" to refer to policies governing the wearing of headscarves, turbans, and chadors. These regulations exist across Europe and several Muslim-majority countries. For example, France banned all types of veiling in state-run institutions, such as schools and hospitals, in 2004. More recently, in 2023, the European Court of Justice (ECJ) ruled that EU Member States may prohibit government employees from wearing religious symbols, including headscarves. Similarly, Kazakhstan, where a majority of women practice Islam, has banned headscarves for students and teachers in schools.

common or socially discouraged. This strengthens the external relevance of our findings by showing that institutional access matters even where social acceptance varies.

In contexts such as Turkey, our findings suggest that integration need not require assimilation into secular norms. Rather, it may involve enabling women's economic and social participation without compelling them to abandon their religious identity. From this perspective, removing identity-based restrictions can promote economic inclusion independently of cultural assimilation. Using data from the Turkish DHS, Uğur (2020) examines the impact of the 1997 ban on wearing headscarves in universities and formal education in a regression discontinuity design by exploiting birth month and year information. She finds no significant effect on university completion among veiled women, largely reflecting their already low university enrollment. Aksoy and Gambetta (2021) combine individual-level survey data on women's veiling from the Turkish Demographic and Health Surveys with regional variation from the 2004 municipal elections, where the Islamic Justice and Development Party (AKP) narrowly won or lost control in different provinces. They find no immediate effect on veiling in 2008 but a significant rise by 2013 in provinces where the AKP narrowly won, particularly among less religious and poorer women, suggesting that veiling can function strategically or instrumentally as the political and social environment becomes more permissive.[5] More closely related to our analysis, Corekcioglu (2021) studies the 2013 removal of the headscarf ban in public spaces using a difference-in-discontinuities design focused on municipal employment outcomes in closely contested mayoral elections. Consistent with our results, she finds an increase in female employment, but only in municipalities governed by Islamist mayors. Our study substantially extends this evidence by examining the entire labor market, covering both public and private sector employment, and showing that the observed gains among veiled women are not fully driven by local governance. In addition, we explore marriage market responses to the repeal and provide suggestive evidence that the removal of the headscarf ban reduced the probability of being married and increased the probability of divorce among veiled women relative to their non-veiled peers.

More broadly, our study contributes to research on how legal and institutional barriers shape women's labor market participation. Recent literature highlights the role of discriminatory laws in limiting women's economic opportunities globally, demonstrating that legal reforms promoting gender equality can significantly boost female employment rates (Roy, 2019; Hyland et al., 2020). Studies of specific institutional reforms also show that removing discriminatory constraints, such as restrictions on women's work hours or barriers in family law, leads to increased female participation in the labor market (Hallward-Driemeier and Gajigo, 2015; Gonzales et al., 2015). More recently, Abou Daher et al. (2025) conduct a field experiment evaluating the effects of lifting Saudi Arabia's ban on women driving and find that treated women are significantly more likely to be employed, though are less able to make purchases without family permission. By framing the headscarf ban as a gender-discriminatory policy, our study extends this literature, demonstrating how removing legal constraints related to religious identity affects pious women's labor market participation.

Finally, our study contributes to the literature on traditional gender norms as barriers to female employment, particularly in developing countries. Jayachandran (2021) emphasizes that restrictive norms concerning women's mobility, household responsibilities, and community expectations significantly reduce female labor force participation. Prior studies have shown limited impacts of economic empowerment policies, such as vocational training or increased control over earnings, due to deeply entrenched gender norms (Field et al., 2010;

[5] Their analysis also examines the immediate aftermath of the 2013 repeal, using veiling information from surveys conducted within two months of the reform. Our study instead focuses on women's labor market responses in the years following the repeal.

Duflo, 2012; Field et al., 2021; Gazeaud et al., 2023). Although veiled women in Turkey generally adhere to conservative gender norms, which are typically associated with lower employment rates (Dildar, 2015), we find that removing the institutional barrier to veiling significantly increased their employment, especially within the public sector. Our findings thus suggest that institutional constraints may substantially outweigh traditional norms in determining women's economic choices, challenging assumptions that conservative norms alone drive low labor force participation among veiled women.

The remainder of the paper is structured as follows: Section 2 discusses the headscarf debate and public sector employment in Turkey. Section 3 describes the data, and Section 4 presents the empirical strategy. Section 5 presents the main findings and robustness checks. Section 6 concludes.

2 Institutional Background

2.1 Historical and social context of veiling in Turkey

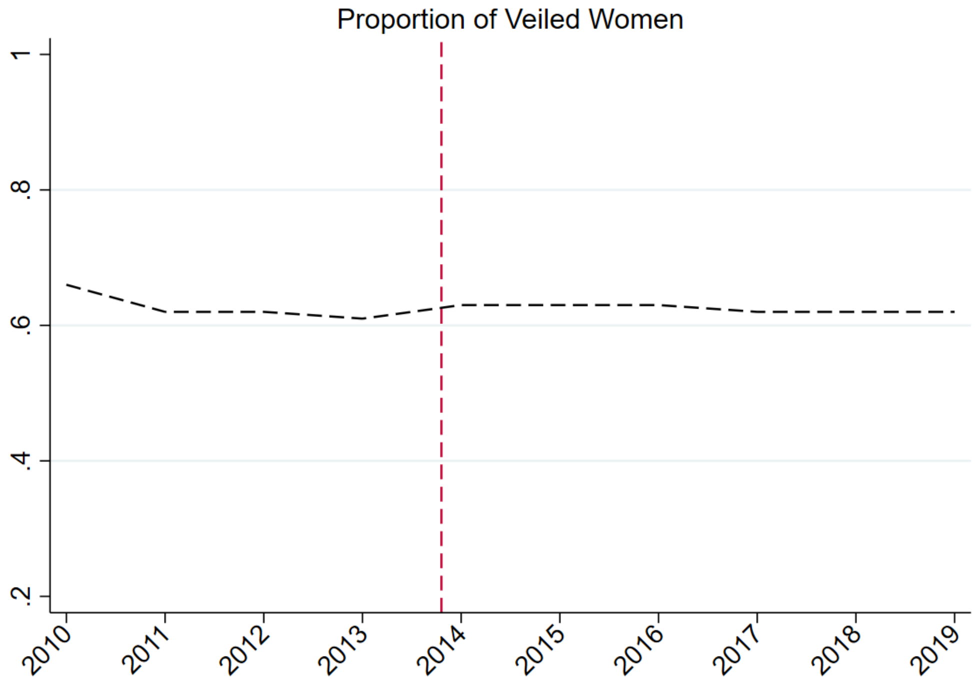
Turkey, constitutionally secular since 1924, has a predominantly Muslim population. The wearing of headscarves in public institutions has long been a contentious issue, contributing to tensions between secularism and religious practices. The historical evolution of headscarf bans in Turkey is examined in detail by Cindoglu and Zencirci (2008). While precise panel data on the prevalence of veiling are limited, multiple sources provide estimates from different time periods. According to the report by KONDA (2019), 97% of Turkey's population identifies as Muslim. Figure 1 illustrates trends in the proportion of women wearing headscarves in Turkey from 2010 to 2019, based on data from these sources. The percentage of veiled women was approximately 66% in 2010, decreasing to 62% in 2011. This figure remained stable until 2013, after which a slight increase occurred (around 2%). Similarly, data from the 2008 and 2013 waves of the Turkish Demographic and Health Survey (DHS) show comparable patterns: 72% of ever-married women reported wearing headscarves in 2008, while 68% of all women did so in 2013. Because the 2008 survey covers only ever-married women, the two figures are not strictly comparable; for ever-married women in the 2013 DHS, the rate is about 74.5%.

Following the 1980 military coup, the first formal headscarf ban was enacted in 1982, which prohibited headscarves in public spaces. This regulation sought to reinforce secularism, a foundational principle of the Turkish Republic. However, the ban sparked significant societal discord, especially among conservative groups who viewed it as an infringement on religious identity and emblematic of secularist authoritarianism. In 1997, the government extended the ban from universities to preparatory schools (Uğur, 2020).

In 2008, the Turkish government made its first significant effort to address the longstanding headscarf ban. This initiative aimed to lift the ban in universities by amending Articles 10 and 42 of the Turkish Constitution. However, the Constitutional Court intervened, ruling on June 5, 2008, that the proposed amendments violated the preamble of the Constitution and the principle of secularism enshrined within it. The Court subsequently annulled the proposed changes, effectively blocking their implementation (Wiltse, 2008; Höjelid, 2010). Despite this setback, the Higher Education Council partially lifted the headscarf ban in 2010 by revising regulations within its jurisdiction. These regulatory changes marked a gradual easing of restrictions on veiling in universities across Turkey (Bianet, 2010; BBC News, 2010).[6]

[6] We discuss in the following section how these regulatory changes may affect our empirical identification.

Figure 1: Share of veiled women in Turkey, 2010–2019



Notes: Data are from KONDA survey data (KONDA, 2019).

A landmark change occurred on October 8, 2013, when the government, under Prime Minister Erdogan, implemented a "democratization package". This reform package aimed to address various human rights and political issues, including improvements in minority rights and democratic governance, such as allowing the use of Kurdish in public schools and enabling the establishment of educational institutions where Kurdish could be taught. While the central focus of this reform package was to expand the rights of minorities, one of its most notable aspects was the dismantling of the longstanding headscarf ban within public spaces. The reform officially came into effect on October 8, 2013, repealing the 1982 by-law that prohibited headscarves.

Prior to this policy change, veiled women were formally prohibited from working in public institutions while wearing a headscarf. Although a small number of veiled women managed to circumvent the ban by wearing wigs in the workplace, this practice remained limited and was not officially recognized by authorities (Cindoğlu, 2011; Guveli, 2011). With the reform, women were legally permitted to wear headscarves in public institutions and government offices, with exceptions for personnel in the armed forces, security forces, and judiciary (Akoglu, 2015; Karahan and Tuğsuz, 2022). The ban was subsequently lifted for members of the judiciary in 2015 and for security forces in 2016.

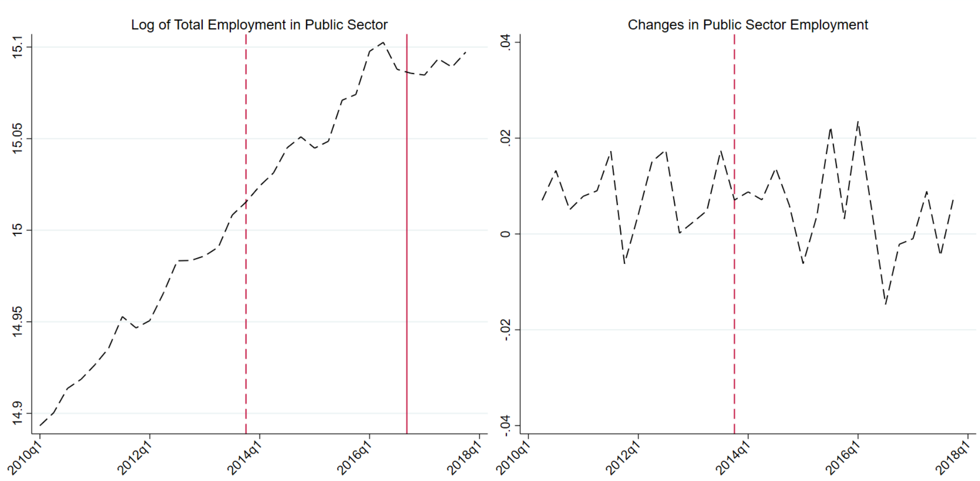
While the repeal of the headscarf ban was part of the Justice and Development Party (AKP)'s broader political agenda and its general direction was therefore predictable (see Corekcioglu, 2021), the precise timing of its implementation was plausibly exogenous to women's labor market choices. Access to public sector employment typically requires passing competitive national examinations, which limits the ability of potential workers to adjust their behavior in anticipation of the policy change. Consistent with this, we show in the next section that pre-repeal trends in public sector employment were similar between treatment and control groups. We also demonstrate that our results are not driven by municipalities governed by AKP mayors, supporting our identification strategy.

2.2 Public sector employment

As of 2023, public sector employment in Turkey accounted for 13% of total employment, below the OECD average of 18%, indicating a relatively smaller public sector presence in the labor market (OECD, 2025). Figure 2 presents trends in total public sector employment in Turkey between 2010 and 2018, based on quarterly data from the Turkish Presidency of Strategy and Budget. The left panel of Figure 2 shows the logarithmic growth in public sector employment, which increased steadily between 2010 and 2016. The upward trend paused around mid-2016, coinciding with the July military coup attempt and a temporary hiring freeze that followed, as the government implemented extensive security reviews and purges (BBC News, 2017; Euronews, 2017). Public sector employment followed a moderate upward trajectory after this period. The right-hand graph of Figure 2 shows that the labor demand side of the public sector remained stable from 2010 to 2017, with no major disruptions apart from the temporary hiring freeze in 2016.

To understand whether the reform was anticipated, it is important to examine the centralized nature of recruitment and exam participation trends. Public sector recruitment in Turkey is centralized and highly standardized. Applicants must pass a biennial civil service examination administered by the Department of Measuring, Selection, and Placement (OSYM). The exam targets different candidate groups in alternating years: odd-year exams primarily recruit for positions such as teachers and district governors, while even-year exams cover a broader range of occupations, including engineers, technicians, and clerks. Figure 3 shows the share of individuals taking the exam during even years between 2008 and 2018. The absence of a noticeable pre-reform increase in participation supports the assumption that the repeal of the headscarf ban was largely unanticipated, though exam participation remains only an indirect proxy for expectations.

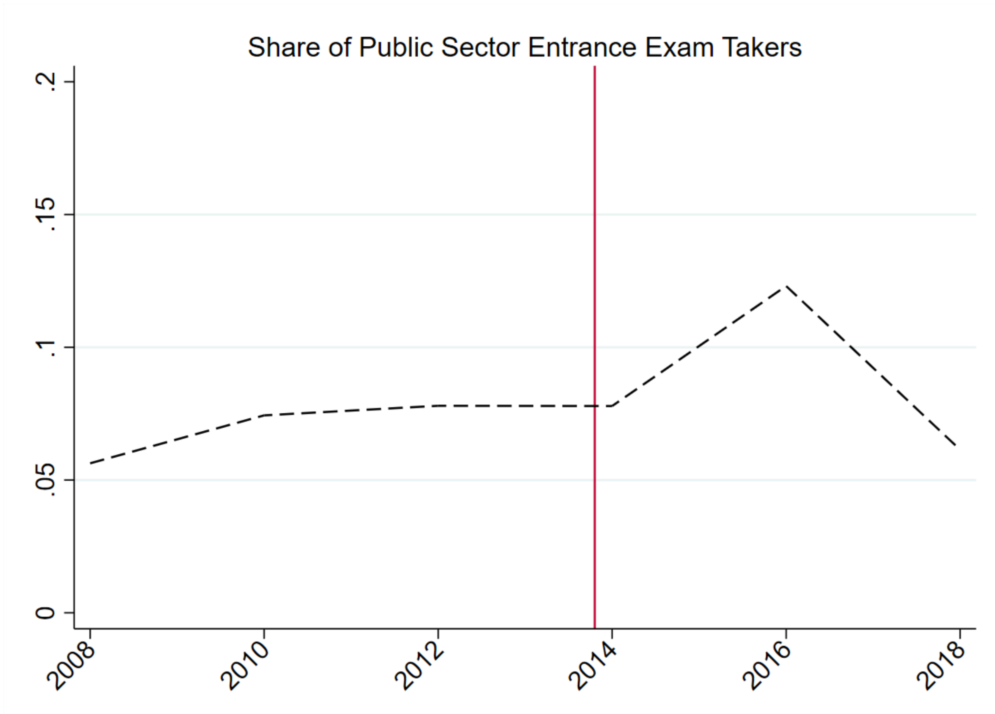
Figure 2: Employment in the public sector as a percentage of total employment



Notes: Data are from the Turkish Presidency of Strategy and Budget.

Public sector positions are predominantly professional roles, such as teaching, nursing, and engineering, that require advanced qualifications. Women hold about 57% of professional public sector positions, compared with 41% among men. Clerical positions, requiring lower educational credentials, employ 17% of women and 15% of men. Technical roles, involving specialized tasks, are the third most common for both genders. In contrast, labor-intensive occupations, including service workers, agricultural and fishery workers, craft and trade workers, and machine operators, are male-dominated, accounting for 28% of male employment versus 9% of female employment. Positions as legislators and senior officials remain uncommon, comprising 4% of male and 3% of female employment. Detailed gender breakdowns across public sector occupations are provided in Table A1 in Online Appendix.

Figure 3: Proportion of individuals taking the public sector entrance exam



Notes: Data are from the Department of Measuring, Selection, and Placement (OSYM).

3 Data

Turkish Household Labor Force Survey. The main dataset used in this study is the Turkish Household Labor Force Survey (HLFS), an annual nationally representative survey conducted by the Turkish Statistical Institute. The HLFS covers around 150,000 households per wave and provides detailed background and labor market information for each household member aged 16 and above. For the analysis, we use repeated cross-sections spanning 2010 to 2017.[7]

The HLFS includes a rich set of variables such as individual characteristics (e.g., age, marital status, education, household size, and region of residence) and detailed labor market information, containing employment type and status (e.g., public or private sector, self-employment or unpaid family work), occupation, weekly hours worked, and monthly earnings. Earnings data incorporate bonuses and premiums received during the reference month.[8]

The analysis focuses on women aged 18–49, who are most likely to participate in the labor market. Because the HLFS reports earnings only for wage earners, analyses of earnings are restricted to this subgroup. In total, the analytical sample comprises 933,810 women. Column 1 of Table A2 in Online Appendix presents descriptive statistics for this sample. Panel A shows that the average age is approximately 33 years, with 98% of respondents being native-born. About 35% have completed at least high school, while 15% hold a university

[7]Data from 2018 onward are excluded to avoid confounding effects from the prolonged political and economic disruptions that followed the July 2016 military coup attempt, including large-scale public sector dismissals. The choice of 2010 as a starting point is motivated by the constitutional referendum held in September 2010, which marked a shift in Turkey's political and institutional trajectory. While the referendum did not directly alter headscarf regulations, it could have shaped expectations about future reforms. By starting in 2010, we ensure that our pre-reform period is defined consistently within the new institutional environment. A remaining concern is that the referendum may have affected treatment and control groups differentially. To address this, we allow for group-specific linear trends in our estimates and show that our main results are robust to this more flexible specification.

[8]Monthly wages are deflated using the Consumer Price Index (CPI) and converted to 2017 Turkish Lira, then expressed in U.S. dollars using 2017 exchange rates.

degree. Roughly half have only a primary school degree (five years of education) or less. Around 71% are married, and the average household size is four persons. Panel B reports labor-market outcomes. About 33% of women are employed, primarily as wage earners (22%), including 5% in the public sector and 17% in the private sector. In addition, 8% work as unpaid family workers, and 3% are self-employed.[9] The remaining columns of Table A2 in Online Appendix present summary statistics separately for women predicted to be veiled and those predicted not to be. In the following section, we discuss the comparability of these samples with those in the DHS, which includes direct information on headscarf use.

Turkish Demographic and Health Survey. While the HLFS provides rich labor market data, the 2013 Turkish Demographic and Health Survey (DHS) uniquely includes information on headscarf use, enabling us to distinguish between veiled and non-veiled women.[10] The 2013 DHS, conducted by Hacettepe University's Institute of Population Studies, is nationally representative and includes 9,746 women, 6,835 of whom were married at the time of the survey (administered between September and December 2013). In addition to headscarf-wearing and other religious practices (e.g., prayer and fasting), the survey contains detailed background and socioeconomic characteristics.

In the 2013 DHS sample, 68% of women aged 18–49 reported wearing a headscarf.[11] Table A3 in Online Appendix compares non-veiled (column 1) and veiled (column 2) women. Non-veiled women are significantly more likely to have completed high school or university, whereas veiled women are concentrated among those with primary education or less. Among non-veiled women, 39% hold a university degree compared with only 6% of veiled women, while 65% of veiled women have primary education or less, compared with 20% of non-veiled women. These differences indicate strong educational sorting by veiling status, with veiled women being substantially less educated on average.

Panel B reports summary statistics for labor market outcomes. Compared with the pronounced education gap between the two groups, the difference in employment probabilities is smaller—around 10 percentage points—but still statistically and economically meaningful, with non-veiled women much more likely to be employed. In addition, only about 25% of veiled women work as wage earners, compared with a substantially higher share among non-veiled women. public sector employment is particularly limited: only about 1% of veiled women hold public sector jobs, compared with roughly 10% of their non-veiled peers, consistent with the restrictions in place before 2013.[12]

4 Empirical strategy

We estimate the effect of lifting the headscarf ban in public institutions on women's labor market outcomes by exploiting both the timing of the policy change and variation in exposure between veiled and non-veiled women within a difference-in-differences (DiD) framework. Using out-of-sample prediction and statistical matching models based on a second dataset (described below), we compare changes in labor market outcomes for women predicted to be veiled with those predicted not to veil, before and after the repeal of the headscarf ban. Formally, we estimate the following equation:

$$y_{it} = \beta_1 + \beta_2\,(Veiled_i \times Post_t) + \beta_3 Veiled_i + X_{it}'\gamma + \mu_t + \varepsilon_{it}. \tag{1}$$

[9] The self-employment category includes women classified as self-employed or own-account workers.

[10] To our knowledge, only the 2008 and 2013 waves of the Turkish DHS collect information on veiling. The 2008 wave covers only ever-married women, whereas the 2013 DHS includes all women aged 15–49.

[11] This figure reflects the unweighted mean; the weighted mean corresponds to 62%.

[12] The presence of veiled women in public sector employment before the reform likely reflects cases where women circumvented the ban by wearing wigs in the workplace, as discussed in Section 2.

where y_{it} represents the outcome of interest for woman i at time t. The indicator $Post_t$ equals one for the post-policy period ($t \geq 2014$) and zero otherwise.[13] X_{it} includes region dummies, age and age squared, household size, and an indicator for being native. The specification also controls for year and region–year fixed effects, which absorb all observable and unobservable shocks common to women within the same region and year.

In a standard DiD framework, the interaction term $Veiled_i \times Post_t$ captures the average treatment effect on the treated. In our context, however, $Veiled_i$ is a predicted rather than an observed indicator, so β_2 represents an intent-to-treat (ITT) effect, capturing the average impact of the reform on women predicted to be veiled relative to those predicted not to be.

Defining treatment and control groups. Our empirical strategy defines treatment and comparison groups based on women's exposure to the removal of the headscarf ban. If veiling status were directly observable in the main dataset, we could estimate the average treatment effect by comparing veiled to non-veiled women. However, such information is not available in large-scale surveys. Because the HLFS does not record veiling, we impute veiling status using an out-of-sample prediction and statistical matching model based on the framework of Rosenbaum and Rubin (1983), estimated with data from the 2013 Turkish Demographic and Health Survey (DHS), which provides nationally representative information on veiling.

The estimation proceeds in two stages. In the first stage, using data from the 2013 DHS,[14] we estimate the probability of wearing a headscarf by fitting a Probit model. The dependent variable equals one if a woman wears a headscarf and zero otherwise. The regressors include education-level indicators, an indicator for being native, age, household size, and region dummies at the NUTS-2 level. Table 1 reports the Probit estimates and marginal effects of each covariate on the probability of veiling. Using these coefficients, we predict headscarf-wearing probabilities for both the DHS and HLFS samples. Although model fit could be improved with additional covariates, the HLFS lacks relevant information beyond these controls.[15]

Figure 4 illustrates the distribution of predicted probabilities of veiling. The left panel displays the distribution based on the DHS sample, and the right panel shows the corresponding distribution for women in the HLFS. The substantial overlap between the two distributions indicates that each woman in the HLFS has comparable counterparts in the DHS based on observable characteristics.

In the second stage, we impute a binary veiling status, equal to one for veiled and zero for non-veiled women in the HLFS, by matching each HLFS observation to its nearest neighbors in the DHS sample based on predicted probabilities.[16] The matched HLFS sample closely mirrors the DHS in population-weighted averages: approximately 61 percent of women in the HLFS are predicted to be veiled (s.d. 0.48), nearly identical to the DHS figure of 62 percent.[17] Columns 2 and 3 of Table A2 in Online Appendix presents descriptive statistics

[13]The headscarf ban in public institutions was lifted in October 2013. Because our data are annual and the survey month is not reported, we designate 2014 as the start of the post-treatment period in the main analysis. As a robustness check, we also re-estimate the specifications treating 2013 as the first post-policy year.

[14]In our main analysis, we estimate veiling probabilities using the 2013 DHS. Although we do not expect major changes in veiling patterns after the policy change, relying on the 2013 data could, in principle, introduce bias if the reform affected veiling choices. To address this concern, we re-estimate veiling probabilities using data from the 2008 DHS wave as a robustness check.

[15]Marital status is excluded to avoid potential endogeneity, as it is analyzed later as an outcome variable. Estimates that include marital status yield results consistent with our baseline.

[16]Matching is performed with replacement. Each DHS observation can be used as a match for multiple HLFS observations. We set the maximum allowed distance in predicted probabilities to 0.05; observations exceeding this threshold are dropped. The excluded observations total 1,416 (less than 1 percent of the estimation sample).

[17]These figures are based on survey sampling weights, which ensure national representativeness. In

Table 1: *Probit estimates of the probability of wearing a headscarf*

Variables	Coef. (1)	Std. Err. (2)	Marginal Eff. (3)
Age	-0.010***	0.003	-0.002***
Native	0.626***	0.135	0.158***
Household size	0.168***	0.017	0.042***
Education level:			
Primary school degree	-0.337***	0.097	-0.068***
Junior high school degree	-0.771***	0.111	-0.186***
High school degree	-1.253***	0.106	-0.347***
University degree	-2.010***	0.114	-0.600***
Obs.	7,172		
Pseudo R^2	0.279		

Notes: Data are from the 2013 Turkish Demographic and Health Survey. The sample includes all women aged 18-49. The estimates include dummies for each region at the NUTS2 level. The omitted category for education is having no formal education. ***, significant at the 1 percent level; **, significant at the 5 percent level; *, significant at the 10 percent level.

by predicted veiling status. For comparison, Table A3 reports the same statistics for the DHS, where veiling is directly observed. The relative gaps between predicted veiled and non-veiled women in the HLFS closely resemble those between observed groups in the DHS, both in background characteristics and in labor market outcomes. This similarity indicates that the predicted veiling status in the HLFS successfully captures the same background differences observed between veiled and non-veiled women in the DHS, suggesting that our imputation model provides a reliable classification for the analysis.

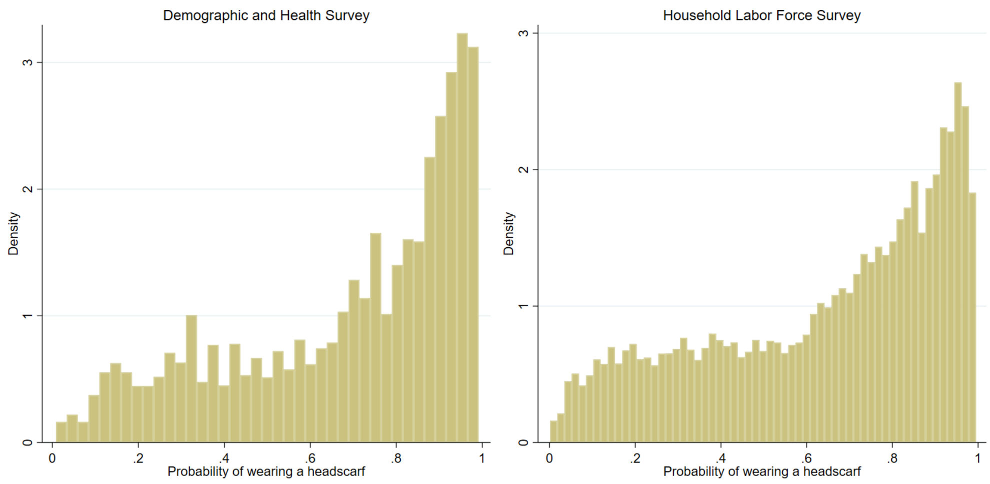
To check how well our model works, we reserve 20% of the DHS, where we actually observe who veils, as a hold-out set: we first fit the model on the remaining 80% of the DHS sample and then evaluate predictions on the hold-out. The model performs well, with a ROC area (AUC) of 0.847 (s.e. 0.010; 95% CI [0.828, 0.867]), meaning a randomly chosen veiled woman is correctly ranked above a randomly chosen non-veiled woman about 85% of the time. We then choose a probability cutoff that balances correct detections and correct rejections by maximizing Youden's J Statistic.[18] At the optimal cutoff (0.63), the model correctly classifies 83% of truly veiled women and 71% of truly non-veiled women in the hold-out sample. Out of roughly 1,070 truly veiled women in the hold-out, the model correctly classifies about 890 (83%) as veiled but misses about 180 (17%), who are misclassified as non-veiled. Similarly, out of roughly 656 truly non-veiled women, it correctly classifies about 465 (71%) as non-veiled but misclassifies about 190 (29%) as veiled. These false negatives and false positives imply that both predicted groups are contaminated: the non-veiled group contains some truly veiled women, while the veiled group contains some truly non-veiled women. As a result, the post-reform averages of the two groups are mechanically pulled toward each other, attenuating the observed difference and leading us to underestimate the true treatment effect, a lower bound. Therefore, our $\hat{\beta}_2$ should be interpreted as intent-to-treat effects.

We verify the robustness of our findings by employing alternative prediction models, including modern machine-learning algorithms. These additional analyses, together with placebo and sensitivity tests, are presented in the next section.

unweighted terms, the corresponding shares are 68 percent in the DHS and 65 percent in the HLFS, reflecting differences in sampling composition rather than population structure.

[18]The resulting threshold is 0.63.

Figure 4: Distribution of predicted probabilities of wearing a headscarf



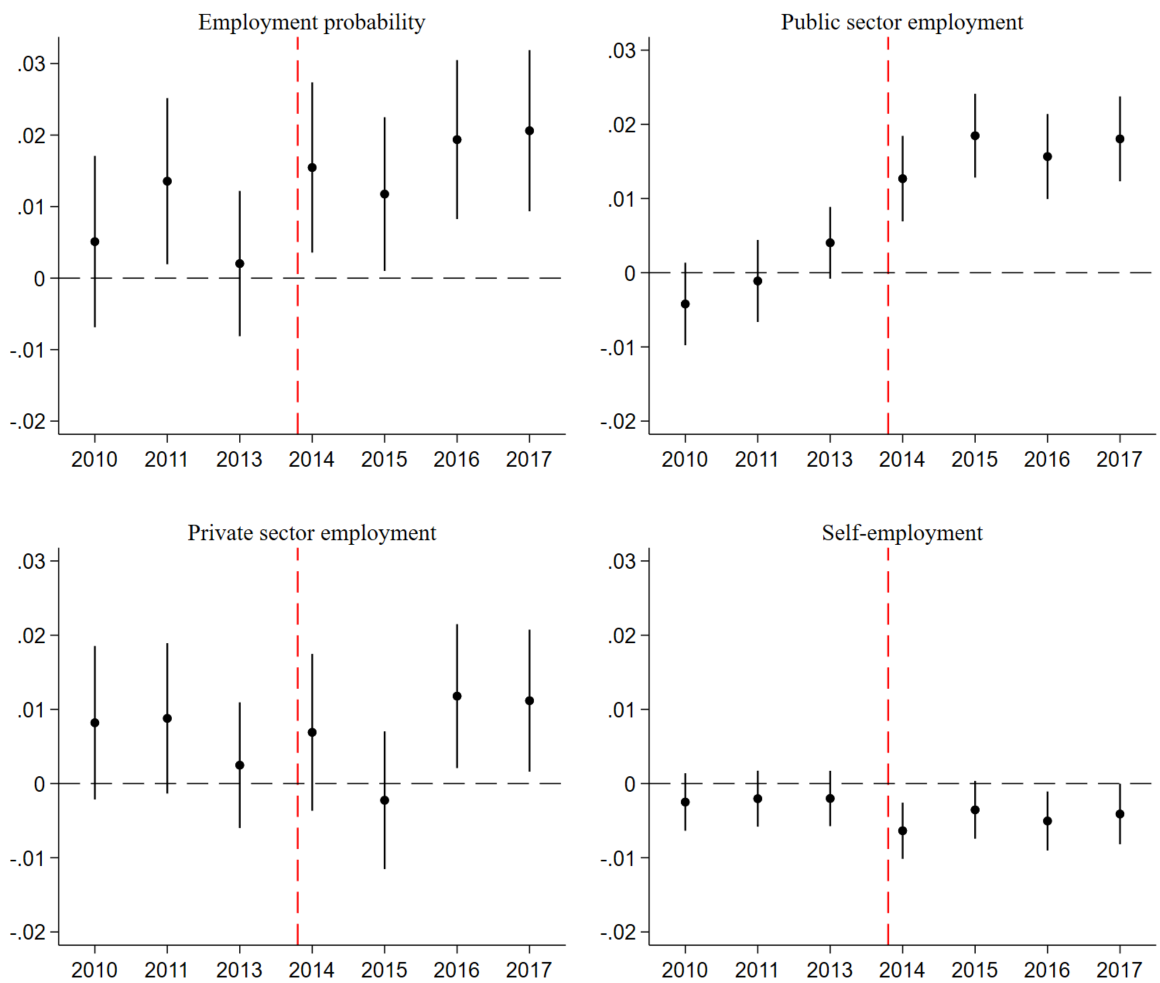
Notes: The left-hand graph shows the distribution of predicted probabilities of veiling based on data from the DHS, while the right-hand graph presents the distribution using data from the HLFS.

4.1 Identifying assumptions

The credibility of our empirical strategy rests on two key conditions: first, that our out-of-sample prediction and matching procedure reliably captures true differences between veiled and non-veiled women; and second, that, absent the reform, labor market outcomes for these groups would have followed parallel trends. We assessed the validity of the first condition in the preceding subsection, showing that predicted veiling status provides a close approximation of actual veiling and enables meaningful comparisons between treatment and control groups. The second condition, the standard identifying assumption in difference-in-differences designs, requires that no other time-varying shocks differentially affected the two groups before the policy change. We examine this assumption using an event-study specification that interacts the veiling indicator with relative-year dummies, taking the last full pre-reform year (2012) as the reference period and controlling for covariates.

The coefficients plotted in Figure 5 show the annual differences between veiled and non-veiled women in the probability of being employed and of working in the public sector, private sector, or self-employment, each measured relative to the 2012 gap, with 95 percent confidence intervals. The results for public and private sector employment, as well as self-employment, illustrate no statistically significant pre-treatment differences, supporting the parallel-trends assumption. For public sector employment, as our data are annual, we observe a modest increase in 2013 relative to 2012, which may reflect that some veiled women began entering public employment soon after the October 2013 policy change, even though 2013 is coded as a pre-treatment year in our main analysis. To account for this possibility, we also re-estimate our models excluding 2013 from the sample as a robustness check. Consistently, joint F-tests of the 2010 and 2011 interaction terms yield $p = 0.23$, $p = 0.21$ and $p = 0.41$ for public, private, and self-employment, respectively. For the probability of overall employment, the pre-trend is marginally significant ($F(2, 31961) = 3.09, p = 0.045$), indicating a modest upward movement between 2011 and 2012. To address this, we estimate a trend-adjusted version of our baseline specification that allows the treatment group its own linear time trend by interacting the veiling indicator with a continuous year variable. In this specification, the post-reform coefficient β_2 captures the change in the veiled–non-veiled gap in the post period, net of the veiled group's own pre-trend. The trend-adjusted estimates are nearly identical to the baseline results, indicating that our conclusions are not driven by the mild pre-period differences visible in 2011. In the following section, we further test the sensitivity of our findings to deviations from the parallel-trends assumption using the framework proposed by Rambachan and Roth (2023).

Figure 5: Trends in labor market outcomes between treatment and control groups



Notes: Data are from the 2010-2017 HLFS. OLS coefficient estimates and their 95% confidence intervals are reported. The controls include whether the woman is a native, dummies for each education level, age, age squared, household size, and region and year fixed effects, with their interactions. Standard errors are clustered at bins of the predicted veiling probability for each woman. Data are weighted using the cross-sectional weights for the wave at which the outcome was measured.

In addition to examining pre-reform trends, it is important to account for other policy changes that may have independently influenced women's labor market outcomes. As discussed in Section 2, the gradual lifting of the headscarf ban in universities beginning in 2010 may have affected women's educational trajectories, particularly by shaping decisions to pursue higher education. Moreover, Turkey implemented two major compulsory schooling reforms known to influence women's education and labor market outcomes. The 1997 reform required individuals born after 1986 to complete junior high school (eight years of schooling), implying that the oldest affected individuals were 30 years old in our sample. A later reform in 2012 extended compulsory education to high school (12 years), but the oldest affected individuals were only 19 by 2017, so only a small share of our sample was exposed. To mitigate confounding from these reforms, we estimate our main specification by including cohort-based exposure indicators that flag (i) eligibility for the gradual lifting of the university headscarf ban beginning in 2010 and (ii) exposure to the 1997 compulsory-schooling reform as robustness checks.

Our identification also assumes that veiling is a relatively stable individual trait and exogenous to the repeal of the headscarf ban over our study period. Sociological and anthropological research emphasizes that veiling is not a transient behavioral choice but a deeply embedded moral and social practice, expressing enduring values of religious identity (e.g., Göle, 1996; Alvi, 2013). These values are cultivated through long-term socialization within families and communities and tend to change only slowly, if at all, in response to short-term political or institutional shifts (e.g., Fleischmann and Phalet, 2012; Drouhot, 2021). While recent work highlights that veiling may also serve instrumental or political purposes (e.g., Shofia, 2022; Aksoy and Gambetta, 2021), such adaptations typically emerge gradually as new social or political equilibria form. We further discuss this assumption in the following section, providing evidence on regional stability in veiling and on how other policy changes that might influence veiling decisions affect our results. Even if the 2013 repeal encouraged some women to begin veiling, whether as an expression of identity or for strategic reasons, such switching would blur treatment and comparison groups and attenuate the DiD contrast, making our estimates conservative lower bounds.

A remaining concern is potential selection on unobservables, particularly religiosity. Because religiosity is positively correlated with veiling and negatively correlated with labor market participation, any omitted-variable bias is likely to attenuate our estimates toward zero. We discuss this issue in detail in the following section, where we present additional analyses designed to account for differences in unobserved religiosity among treated women.

5 Main Results

5.1 Effects on Labor Market Outcomes of Women

We begin by presenting the intent-to-treat (ITT) estimates of the effect of lifting the headscarf ban on women's labor market outcomes. Panel A of Table 2 reports the estimated coefficients derived from Eq. (1), while Panel B allows for group-specific pre-trends by interacting the veiled indicator with a linear year trend, as discussed in Section 4.1. Our focus is on the interaction term $Veiled \times Post$, which captures the relative change in outcomes for women predicted to be veiled compared to those predicted not to veil after the reform.

Column 1 shows that the policy led to a statistically significant increase in the employment probability of veiled women relative to their non-veiled peers, by about 1.2 percentage points. This corresponds to an increase of roughly 4–5 percent relative to the pre-policy mean for veiled women. Trend-adjusted estimates in Panel B yield very similar results, though the employment effect becomes only marginally significant, as expected given the added flexibility in this specification. Columns 2 and 3 decompose wage employment into public and private sector components. The strongest response is observed for public sector employment: the probability that veiled women worked in the public sector rose by 1.7 percentage points relative to their non-veiled peers. Given that the pre-reform mean among veiled women was only 1.7 percent, this effect is substantial, roughly doubling their pre-reform rate, as reported in the last row of Table 2. The increase remains positive and statistically significant under the trend-adjusted specification, although its magnitude falls slightly to 0.9 percentage points. By contrast, the coefficient on private sector employment is small and statistically insignificant, suggesting that the reform's immediate effects were concentrated within the public sector rather than the private sector.

Columns 4 and 5 present results for self-employment and unpaid family work, respectively. The estimated coefficients reveal a decline in both outcomes: self-employment fell by about 0.3 percentage points in Panel A (and 0.5 percentage points in Panel B), while unpaid family work declined by roughly 1.0 percentage point in Panel A, although the estimate becomes less precise once group-specific trends are included.

Overall, the findings indicate that the repeal of the headscarf ban not only increased employment among veiled women but also changed the sectoral allocation of their work. The sharp rise in public sector employment, coupled with declines in self-employment and unpaid family work, suggests that some veiled women who were already in the labor market reallocated from informal or unpaid jobs toward formal public positions once these became accessible.

Table 2: *Effects of the headscarf ban removal on women's labor market outcomes*

Outcome	Employment probability (1)	Public sector employment (2)	Private sector employment (3)	Self-employed (4)	Unpaid family worker (5)
Panel A: Main estimates					
Veiled × Post	0.012***	0.017***	0.002	-0.003***	-0.010***
	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)
Veiled	-0.010***	-0.006***	-0.003	0.000	0.001
	(0.003)	(0.002)	(0.003)	(0.001)	(0.002)
Panel B: Trend-adjusted estimates					
Veiled × Post	0.011*	0.009***	0.002	-0.005**	-0.003
	(0.006)	(0.003)	(0.006)	(0.002)	(0.003)
Veiled	-0.010**	-0.009***	-0.003	-0.000	0.004*
	(0.004)	(0.002)	(0.004)	(0.001)	(0.002)
Observations	933,810	933,810	933,810	933,810	933,810
Mean Dep. Var.	0.275	0.017	0.116	0.039	0.110

Notes: Data are from the 2010-2017 HLFS. The sample includes all women aged 18-49. Panel A reports difference-in-differences estimates comparing labor market outcomes between predicted to be veiled and non-veiled women. Panel B includes $1\{\text{Veiled}\} \times t$, where t is a centered year variable (t = year − 2010). Each estimate includes a dummy variable indicating whether the woman is native, dummies for the woman's education level, age, age squared, household size, and region and year fixed effects, along with their interactions. Standard errors (in parentheses) are clustered at bins of the predicted veiling probabilities for each woman. The last row reports the mean outcome for the sample of veiled women before the policy change. All estimates are weighted using cross-sectional survey weights from the wave in which the outcome was measured. ***, significant at the 1 percent level; **, significant at the 5 percent level; *, significant at the 10 percent level.

To better understand how the removal of the headscarf ban affected women's labor market outcomes, we next examine weekly working hours, where non-employed women are coded as working zero hours. Column 1 of Table 3 shows that lifting the headscarf ban increased veiled women's weekly working hours by about 0.5 hours, corresponding to a 4 percent rise relative to the pre-policy mean for veiled women. The effect remains positive and statistically significant under the trend-adjusted specification in Panel B, with an estimated increase of 0.7 hours. This increase primarily reflects higher employment among veiled women rather than longer hours among those already employed. If the policy had affected hours at the intensive margin, we would expect a larger increase in average weekly hours. Instead, the magnitude aligns closely with the rise in employment probability reported in Table 2, indicating that the reform operated mainly through new employment opportunities rather than increased working hours for veiled women who were already employed before the repeal of the ban.

Column 2 of Table 3 examines monthly labor earnings, measured only for employed wage earners.[19] The estimates show no robust effect: panel A suggests a modest positive coefficient—roughly 6 percent of the pre-policy mean, but it is imprecisely estimated and becomes insignificant once group-specific pre-trends are allowed in Panel B. This result

[19]Earnings are available only for wage earners, so the analysis excludes the self-employed and unpaid family workers.

should be interpreted cautiously, as post-reform selection into employment likely biases the estimates, limiting causal interpretation.

Columns 3 and 4 turn to indicators of job quality. Column 3 shows that the reform reduced the probability that veiled women worked without social security coverage by about 0.9 percentage points, consistent with a shift from informal to formal public sector jobs. Column 4 shows a corresponding improvement in contract stability: the probability of holding a permanent contract increased by 1.8 percentage points (Panel A), equivalent to roughly a 16 percent rise relative to the pre-policy mean. This effect remains positive, though smaller in magnitude, under the trend-adjusted specification in Panel B. These findings are consistent with the observed rise in public sector employment, as such positions in Turkey typically provide formal contracts and social security coverage.

The remaining columns of Table 3 explore the occupational distribution of veiled women within the public sector. Each outcome is a binary indicator equal to one if a woman is employed in a specific public occupation and zero otherwise. Column 5 captures professional occupations requiring higher education (e.g., teachers, doctors, nurses), Column 6 covers technicians and associate professionals, and Column 7 includes clerical occupations such as clinical assistants, midwives, administrative secretaries, and office clerks. The results indicate that veiled women gained access to higher-skill jobs in the public sector. The probability of working in professional roles increased by 0.3–0.6 percentage points, while the likelihood of employment as a technician rose by about 0.9 percentage points in the baseline specification, with a smaller but still positive effect in the trend-adjusted estimates. We also find suggestive, though less robust, evidence of an increase in clerical positions. Overall, these results suggest that the expansion in public sector employment among veiled women was concentrated in professional and technical occupations, reflecting improved access to formal, higher-quality jobs once the ban was lifted.

Finally, to complement our main results, Online Appendix Table A4 presents analogous estimates for women predicted not to be veiled. We re-estimate equation (1) in a simplified form by replacing the year fixed effects with a single post-period indicator. This coefficient captures aggregate differences in outcomes between the pre- and post-policy periods for non-veiled women and should be interpreted as descriptive, reflecting compositional adjustments rather than causal effects of the reform. The estimates show that, after 2013, non-veiled women became less likely to work in the public sector, while their likelihood of private sector employment and self-employment increased. Given that the overall size of public sector employment remained stable during this period (Figure 2), these patterns are consistent with a reallocation in the composition of public sector jobs, with veiled women gaining access once the ban was lifted and non-veiled women correspondingly shifting toward private- or self-employment.

All in all, the evidence paints a clear picture of how the repeal of the headscarf ban reshaped women's labor market opportunities. Veiled women gained access to public sector jobs that had previously been closed to them, leading to higher overall employment. The increase in public sector employment among veiled women was driven not only by higher employment rates but also by a reallocation from self-employment and unpaid family work into formal public sector positions. These new opportunities were concentrated in professional, technical, and clerical occupations—roles typically requiring some educational qualifications, suggesting that better-educated veiled women were the main beneficiaries. Meanwhile, non-veiled women became less likely to hold public sector jobs compared with the pre-policy period, consistent with a compositional shift in which veiled women partly substituted for non-veiled women once the ban was lifted. Overall, the evidence underscores how institu-

Table 3: *Effects of the removal of the headscarf ban on working hours, job quality, and occupations*

Outcome	Working hours (weekly)	Earnings (monthly)	Informal employment	Permanent job	Professionals	Technicians	Clerks
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Main estimates							
Veiled × Post	0.491***	10.695*	-0.009***	0.018***	0.003**	0.009***	0.003***
	(0.148)	(6.161)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
Veiled	-0.388***	-4.336	-0.001	-0.009***	-0.001	-0.004***	-0.001***
	(0.150)	(5.613)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
Panel B: Trend-adjusted estimates							
Veiled × Post	0.685**	10.750	-0.007*	0.011*	0.006**	0.002**	0.000
	(0.311)	(12.435)	(0.004)	(0.006)	(0.003)	(0.001)	(0.001)
Veiled	-0.316	-4.316	-0.001	-0.011***	0.001	-0.006***	-0.002***
	(0.200)	(7.298)	(0.003)	(0.004)	(0.001)	(0.001)	(0.001)
Observations	915,873	790,783	933,810	933,810	933,810	933,810	933,810
Mean Dep. Var.	10.95	180.8	0.181	0.114	0.008	0.004	0.003

Notes: Data are from the 2010-2017 HLFS. The sample includes all women aged 18-49. Panel A reports difference-in-differences estimates comparing labor market outcomes between predicted to be veiled and non-veiled women. Panel B includes $1\{\text{Veiled}\} \times t$, where t is a centered year variable ($t = \text{year} - 2010$). Each estimate includes a dummy variable indicating whether the woman is native, dummies for the woman's education level, age, age squared, household size, and region and year fixed effects, along with their interactions. The working hours and labor market earnings variables take zero for those who are unemployed. The earnings variable is only available for those who are employed as wage earners in the survey. Standard errors (in parentheses) are clustered at bins of the predicted veiling probabilities for each woman. The last row reports the mean outcome for the sample of veiled women before the policy change. All estimates are weighted using cross-sectional survey weights from the wave in which the outcome was measured. ***, significant at the 1 percent level; **, significant at the 5 percent level; *, significant at the 10 percent level.

tional barriers to identity-based expression can shape not only labor market participation but also the distribution of high-quality jobs across groups.

Exogeneity of veiling status. A potential concern for interpreting our results arises if the policy change induced some women to adopt veiling. For instance, a woman who previously refrained from veiling to pursue a career in the public sector might begin veiling and subsequently enter public employment after the ban was lifted. Because our empirical design classifies women into treatment and comparison groups based on pre-policy predictions of veiling status, which we assume to be time-invariant, such switchers would be misclassified. These women, who experience a positive labor market response to the reform, would remain in the comparison group, thereby contaminating its post-policy outcomes. Conversely, some women with low predicted veiling probabilities might begin veiling strategically after the repeal, adopting the headscarf not out of religious conviction but to access newly opened public sector opportunities. Since their pre-policy characteristics place them in the non-veiled (comparison) group, post-policy changes in their labor market outcomes would again raise outcomes in the control group, reducing the estimated treatment effect.

While we cannot directly observe such switching in our main data due to the lack of individual panel information,[20] we present several pieces of evidence suggesting that veiling is highly stable over time and unlikely to respond to short-term policy changes.

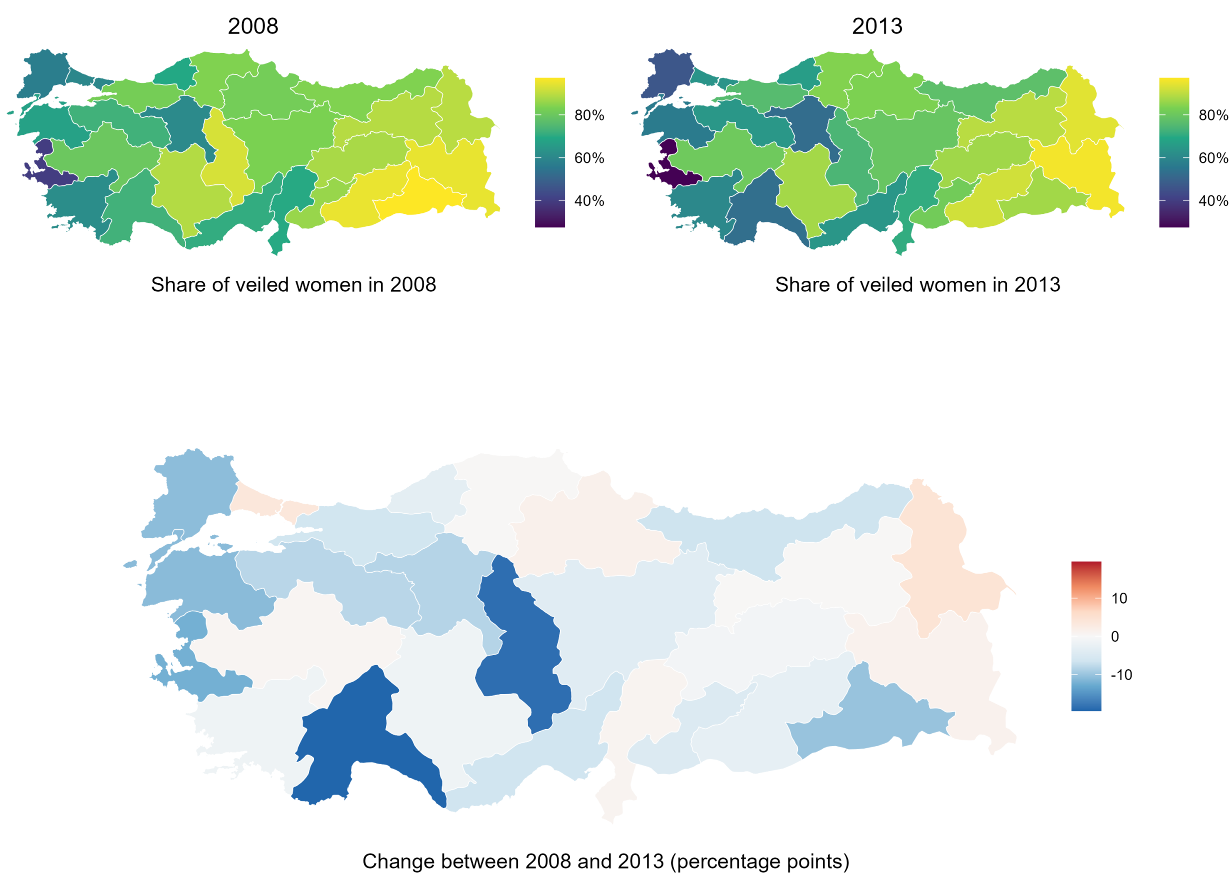
We first assess the persistence of veiling at the aggregate level. Figure 6 plots the regional share of ever-married women aged 18 and above who report wearing a headscarf in 2008 and 2013, using DHS data.[21] The two top panels show the regional shares in 2008 and 2013,

[20] Only the 2008 and 2013 Turkish DHS waves collect information on veiling. To our knowledge, KONDA is the only institution maintaining a panel dataset on veiling, but access is proprietary and requires purchase.

[21] We restrict the sample to ever-married women because the 2008 DHS collects veiling information only for this group.

while the bottom panel presents the change in percentage points between the two surveys. The comparison reveals a striking degree of regional stability: the Pearson correlation in regional veiling rates between 2008 and 2013 is 0.93, and the Spearman rank correlation is 0.89 (both $p < 0.001$). This implies that both absolute levels and the relative ranking of regions remained almost unchanged over the five-year period.

Figure 6: Regional stability in veiling, 2008–2013



Notes: Data are from the 2008 and 2013 Turkish Demographic and Health Survey. This figure plots the regional share of ever-married women aged 18 and above who report wearing a headscarf since the 2008 DHS covers data only for the sample of ever married women. Regions correspond to the 26 NUTS-2 areas of Turkey. The two panels in the top row show the proportion in 2008 and 2013, while the bottom panel shows the change in percentage points.

Second, we examine whether earlier institutional reforms in Turkey affected veiling behavior. The first is the 1997 extension of compulsory schooling from five to eight years, which is well documented to have increased women's educational attainment and labor market participation (Erten and Keskin, 2018; Güneş, 2016; Merlino and Yurdakul, 2024). The second is the 2002 reform of the Civil Code, which eliminated several legal provisions favoring men (e.g., polygyny, unilateral divorce, inheritance rights) and raised the minimum marriage age to 18 for both genders (Anıl, 2002; Kırdar et al., 2018). Using the 2013 DHS and exploiting variation in birth month and year in a regression discontinuity design, we assess both reforms (see Online Appendix Section A and Tables A5–A6). Across specifications, we find no statistically or economically significant effects of either reform on the probability of veiling. We also find no impact on regular prayer, and only a small negative effect of the 2002 reform on fasting during Ramadan.

Taken together, the regional stability and the null effects of earlier institutional reforms suggest that veiling is a relatively persistent practice, shaped more by enduring cultural and religious norms than by short-run policy variation. Nevertheless, to the extent that some switching in response to the headscarf ban may have occurred, such misclassification would attenuate our estimated effects by introducing treated individuals into the comparison group. As a result, our estimates should be interpreted as conservative lower bounds of the true causal impact of the reform.

Unobservable heterogeneity. Another potential concern is that veiling status may be correlated with unobserved religiosity, which can influence labor supply decisions, as discussed by Carvalho (2013). This correlation could bias our estimates in either direction. For instance, if the most religious women, those for whom veiling is non-negotiable, were also the most constrained by the ban, they might respond most strongly to its repeal. In this case, our estimates would capture both the policy effect and a differential response among highly religious women (an upward bias). However, the well-documented negative association between religiosity and female employment (Dildar, 2015) suggests that a substantial upward bias is unlikely.

Because religiosity is unobserved, we proxy for it using regional veiling prevalence prior to the reform. Specifically, we interact our main specification with an indicator for residing in a region with an above-median pre-reform share of veiled women, based on the 2013 DHS.[22] If unobserved religiosity were driving the results, we would expect larger effects in high-veiling regions. Panel A of Table 4 shows no such pattern. The triple-interaction term ($Veiled \times Post \times HighShare$) is statistically insignificant for the probability of employment and significantly negative for public sector employment. While veiled women in high-veiling regions gained about 1.3 percentage points in public sector employment, the effect in low-veiling regions was larger, around 2 percentage points. This pattern contradicts the notion of upward bias from religiosity. If anything, it suggests a potential downward bias: less religious veiled women, who were more likely to be constrained by the ban, appear to have responded more strongly to the new economic opportunities. Combined with the measurement error in predicted veiling status, explained above, which mechanically attenuates treatment effects toward zero, these results indicate that our estimates are conservative, providing lower-bound estimates of the true causal impact of the reform.

Although we assume religiosity to be time-invariant, it may still evolve differently across groups. To strengthen our causal interpretation, we examine heterogeneity by education. To do so, we compare veiled women with no formal education or only primary school degree (5 year of education) to those with at least a junior high school degree ($HigherEduc$). Since the headscarf ban applied only to public institutions, its repeal should primarily benefit relatively more educated veiled women, who are more likely to be eligible for such jobs rather than those with only a primary school degree or no formal education. Panel B of Table 4 confirms this pattern. The triple-interaction coefficients are positive and highly significant for both the probability of employment and public sector employment, indicating that the observed effects of the reform are concentrated mainly among higher-educated veiled women. The reform increases the probability of employment primarily for relatively more educated group (about 1.1 percentage points), whereas the effect for less-educated veiled women is small and statistically insignificant. Similarly, the probability of public sector employment rises for both groups, but the effect is roughly five times larger for more educated peers (1.4 percentage points), than the low-educated veiled women who experience a 0.3 percentage point increase. These results suggest that the removal of the ban provided public sector opportunities primarily to women with some educational qualifications required for such positions. The declines in self-employment and unpaid family work observed in the

[22] Results are similar when using the 2008 DHS.

Table 4: *Heterogeneous effects*

Outcome	Employment probability (1)	Public sector employment (2)	Private sector employment (3)	Self-employed (4)	Unpaid family worker (5)
Panel A: Regional share of veiled women					
Veiled × Post × HighShare	0.001	-0.007**	0.005	-0.007***	0.004
	(0.006)	(0.003)	(0.005)	(0.002)	(0.003)
Veiled × Post	0.011**	0.020***	-0.001	-0.001	-0.012***
	(0.005)	(0.002)	(0.004)	(0.002)	(0.002)
Panel B: Education level					
Veiled × Post × HigherEduc	0.025***	0.011***	0.004	0.006**	0.007*
	(0.006)	(0.002)	(0.005)	(0.002)	(0.004)
Veiled × Post	-0.014***	0.003***	-0.006	-0.005**	-0.009***
	(0.005)	(0.001)	(0.004)	(0.002)	(0.003)
Observations	933,810	933,810	933,810	933,810	933,810

Notes: Data are drawn from the 2010–2017 HLFS and the sample includes all women aged 18–49. Each panel presents heterogeneity analyses by regional share of veiled women, education level, and regional Islamist party vote share, respectively. The table reports difference-in-differences estimates comparing labor market outcomes between women predicted to be veiled and those predicted not to be veiled. All specifications include a dummy variable indicating whether the woman is native, dummies for the woman's education level, age, age squared, household size, and region and year fixed effects, along with their interactions. Standard errors (in parentheses) are clustered at bins of the predicted veiling probabilities for each woman. All estimates are weighted using cross-sectional survey weights from the wave in which the outcome was measured. ***, significant at the 1 percent level; **, significant at the 5 percent level; *, significant at the 10 percent level.

aggregate are driven mainly by low-educated veiled women. In contrast, higher-educated veiled women exhibit net employment gains concentrated in the public sector. If religiosity alone explained labor market entry, we would expect broad increases across education groups, including among low-educated veiled women. Instead, the effects are concentrated exactly where they should be if the ban constrained access to public employment among educated veiled women. This interpretation aligns with our earlier results showing that post-reform public sector employment gains are concentrated in professional and technical occupations requiring formal educational qualifications.

5.2 Robustness checks

Parallel trends. We assess the sensitivity of our main results in Table 2 to remaining pre-trend differences using the falsification-robust procedure of Rambachan and Roth (2023). The method places a relative-magnitude bound M on any post-treatment deviation from parallel trends, scaling it to the largest adjacent pre-period deviation. For each outcome we report a breakdown value $\bar{M}$, defined as the smallest M at which the falsification-robust confidence interval (FLCI) first contains zero. Online Appendix Figure A1 plots the 95% FLCIs together with the baseline point estimates for $M \in \{0, 0.1, \ldots, 0.5\}$. As M increases, the parallel-trends assumption is relaxed and the FLCIs widen. An effect is robust up to $M^\star$ if the FLCI remains strictly above (or below) zero for all $M \leq M^\star$; the breakdown value $\bar{M}$ marks the first relaxation at which the confidence interval touches zero.

The results show a clear pattern of robustness across outcomes. For the employment probability, the effect remains statistically positive up to $\bar{M} \approx 0.4$, meaning it survives any post-policy drift up to 40% as large as the worst pre-policy deviation; beyond this threshold (e.g. $M = 0.5$) the effect is no longer distinguishable from zero. public sector employment is more robust, with $\bar{M} > 0.5$ (no breakdown on our grid), indicating that the positive effect persists even under sizeable deviations from parallel trends. By contrast, private sector employment exhibits $\bar{M} \approx 0$, consistent with no effect. Finally, the negative effects on self-employment and unpaid family work are highly robust, as the FLCIs remain strictly below zero for all $M \leq 0.5$.

Other potential confounding factors. We start by assessing the robustness of our results to the definition of the post-policy period. While the headscarf ban in public institutions was officially lifted in October 2013, our data are annual, making it possible that using 2014 as the post-policy period underestimates the policy's anticipation effect. The 2013 wave might partially reflect early responses to the policy change. To address this, we re-estimate the main results excluding the 2013 wave. The coefficients, reported in Panel A of Table A7 in Online Appendix, remain stable in both magnitude and precision, indicating that the observed effects are not artifacts of timing misclassification. In the context of Turkey, we assess sensitivity of our results to other relevant policy changes that could potentially explain our results rather than the reform. First, as discussed in 2, starting from 2010, universities in Turkey gradually began allowing female students and employees to wear headscarf gradually, with institutional variation in implementation. This could potentially have influenced labor market outcomes, particularly for women born after January 1992, who were more likely to enter university after these changes. To test whether our results are driven by this cohort, we introduce a binary indicator for women born in or after January 1992 and interact it with the *Post* variable. Results in Panel B remain consistent with our main estimates, suggesting that the observed labor market effects are not confounded by earlier initiatives targeting the ban across universities. Second, we control for the potential effects of the 1997 education reform, which extended compulsory schooling from five to eight years for cohorts born after 1986, in which women born in January 1987 or after are assigned as affected cohort and those born before could drop out after 5 years of education. Labor market effects of this reform for women well documented (Erten and Keskin, 2018; Merlino and Yurdakul, 2024; Güneş, 2016). This reform likely affected women's educational attainment and subsequent labor market outcomes. To account for this, we introduce a binary indicator for women born after 1986 and interact it with the post-policy period. As reported in Panel C of Table A7, our main findings remain robust, indicating that the observed effects of the headscarf ban's removal are distinct from those of earlier education reforms.

While the above results support that our findings are primarily attributable to the repeal of the headscarf ban, a remaining concern in the Turkish context is whether the observed gains are partly driven by political favoritism rather than the policy itself. Specifically, if municipalities governed by Islamist parties were more inclined to hire veiled women after the reform, the estimated effects might reflect preferential hiring in politically aligned areas rather than the removal of an institutional barrier (Corekcioglu, 2021). To explore this mechanism, we construct a measure of Islamist political influence based on the vote share of the Justice and Development Party (AKP) in the 2014 municipal elections. Regions with an above-median AKP vote share (46.8 percent) are classified as high political influence, and this indicator is interacted with our difference-in-differences terms. It is important to note, however, that this analysis should not be interpreted as causal heterogeneity. If the reform itself affected subsequent political support for the AKP, political alignment may be endogenous to the policy change. We therefore view this analysis as a descriptive mediation test, assessing whether the labor-market effects in Table 2 are concentrated in AKP-running regions. The results in Panel D of Table A7 show that the reform's effects are not confined to politically aligned regions. Although veiled women in high-AKP regions experience somewhat larger gains, particularly in overall and public sector employment, the effect on public sector employment remains positive and statistically significant even in less Islamist regions. Moreover, we find no differential effects in private sector outcomes. These findings suggest that while Islamist-leaning municipalities may have amplified access for veiled women, the primary driver of the observed gains is the nationwide removal of the institutional barrier that had previously excluded them from public employment.

Alternative out-of-sample prediction models. Our main analysis imputes veiling status using a model trained on the 2013 wave of the Demographic and Health Survey (DHS). To assess whether our results depend on this particular training sample, we re-estimate the main specification using veiling probabilities predicted from the 2008 DHS, which reports veiling information for a representative sample of ever-married women. The estimated treatment effects, reported in Panel E of Online Appendix Table A7, are very similar to our baseline results in both magnitude and significance. This supports the stability of the underlying relationship between individual characteristics and veiling behavior, and indicates that our findings are not sensitive to the choice of DHS wave used for prediction and imputation.

To further validate our results, we explore alternative treatment and control groups. In the first exercise, we redefine our treatment group as women who hold traditional attitudes toward gender roles, those who agree with the statement that *"the husband should work while the wife stays home"* (i.e., traditional). If the estimated treatment effects for women who are predicted to be veiled closely resemble those for traditional women, it would be difficult to determine whether the effect is specific to veiled women or the sample of women with traditional attitudes. In principle, given that the low correlation between being traditional towards gender role attitudes and being veiled (0.20 in the DHS sample), we should not expect similar effects of the reform on these samples. We predict traditional attitudes following the same methodology as our main analysis and re-estimate Eq. (1) using the variable $Traditional$ instead of $Veiled$. Results presented in Panel F indicate smaller and statistically distinct effects for traditionally-minded women. the Wald test of equality of coefficients confirms significant differences in the estimated effects for employment (p-value = 0.000) and public sector employment (p-value = 0.000). These findings support that the effects of the reform are indeed specific to veiled women rather than women with traditional attitudes more generally.

In the second exercise, we use the frequency of praying daily to validate our main results. Given the strong correlation between regular prayer and veiling status (0.42 in the DHS sample), this provides a more behaviorally anchored proxy for veiling. If the estimated effects are similar for women who pray daily, this would further validate our out-of-sample predictions. The results reported in Panel G indicate significant increases in employment probability and public sector employment for women who pray regularly, along with declines in unpaid family work. Importantly, the Wald test for employment (p-value = 0.734) suggests no significant difference between effects estimated for veiled women and those who pray daily, further supporting the validity of our prediction model. However, the effect on public sector employment is larger for veiled women (p-value = 0.000), which is expected given that the reform directly addressed institutional barriers specific to veiling.

To improve the validity of the veiling imputation, we further employ alternative machine learning (ML) techniques to predict headscarf wearing. We then use the ML-predicted veiling information to estimate the policy effects. Overall, the results remain broadly consistent with those obtained using propensity score matching. Machine learning methods accommodate high-dimensional and nonlinear relationships between observables and veiling, thereby strengthening the credibility of the imputation. The procedure involves two steps. First, the DHS sample is divided into a training set (80%) and a test set (20%). The training set is used to estimate a predictive model that links women's background characteristics to veiling status, and the test set is used to evaluate the model's predictive performance. Second, the trained model is applied to the HLFS to generate out-of-sample predictions of veiling.

Our main approach relies on ensemble learning, which combines three models—Support Vector Machines, Random Forest, and K-Nearest Neighbors—using soft voting. Each model is first trained separately to predict the probability that a respondent is veiled. The ensemble model then averages these predicted probabilities across the three algorithms, classifying respondents with an average probability above 0.5 as veiled. This approach integrates information from multiple models and improves predictive accuracy. The trained ensemble achieves approximately 80% accuracy on the DHS training set. When applied to the HLFS, about 66% of women wear a headscarf—closely matching the DHS average. To assess the robustness of our results to the choice of algorithm, we also re-estimate the treatment effects using each model individually—Support Vector Machines, Random Forest, and K-Nearest Neighbors—as well as an alternative Neural Network model.[23] Together, these five models represent the most widely used and credible machine learning approaches for prediction tasks.

Table A8 reports the estimated policy effects on women's labor market outcomes using the ML-imputed data. Consistent with the main analysis, the results show statistically significant increases in employment probability and public sector employment among veiled women following the policy change, accompanied by declines in self-employment and unpaid family work. While the magnitudes and significance levels vary slightly across models due to differences in algorithms, the overall patterns remain stable across panels. Unlike the PSM results, private sector employment shows a statistically significant increase at the 5% level in the Support Vector Machine and ensemble learning models—the latter likely driven by the former. However, since the effect does not appear consistently across the remaining three models, we treat it as suggestive rather than conclusive.

Table A9 examines the policy effects on working hours, job quality, and occupational composition using ML-imputed veiling status, serving as a robustness check to Table 3. The results consistently show positive effects on women's working hours and the likelihood of holding a permanent job, alongside a corresponding decline in informal employment. The estimates also indicate that veiled women are more likely to work in occupations requiring lower educational attainment, such as technician or clerical positions. In contrast, all models show insignificant effects on monthly earnings. Overall, these robustness checks confirm the stability of the labor market effects across different machine learning specifications.

In sum, these robustness exercises indicate that our results are not driven by arbitrary classification rules, omitted variable bias, or concurrent policy shocks. More importantly, the ML-based out-of-sample prediction approaches corroborate our main findings across a range of methodologies, thereby providing a rigorous framework for evaluating policy effects in settings where direct treatment assignment is unobservable.

5.3 Discussion

Our findings indicate that lifting the headscarf ban significantly increased public sector employment among women predicted to be veiled, with the gains concentrated among relatively more educated women who are most likely to qualify for public jobs. This suggests that the ban had previously excluded some of the most productive women from public employment.

[23] The Neural Network produces continuous predicted probabilities rather than discrete classifications. As a result, the reported coefficient reflects the interaction between the predicted probability of veiling and the post-policy indicator.

These results carry important implications for other Muslim-majority contexts where veiling restrictions remain in place. For example, Kazakhstan introduced a ban on wearing hijabs in schools in 2023.[24] While such policies are often justified on secularist or uniformity grounds, our evidence suggests that institutional restrictions may impose substantial productivity costs by limiting the labor market participation of relatively more productive veiled women.

Although our analysis focuses on a Muslim-majority setting, the findings also raise questions for Muslim-minority contexts, such as those in Western Europe. Within Turkey, the effects of the reform do not significantly differ across regions with above- versus below-median veiling rates (Panel A of Table 4). In Online Appendix Table A10, we further compare women who are a local minority (in regions with veiling rates below 50%) to those in regions where veiling is more common, and again find no significant differential effects. These results suggest that the impact of removing institutional barriers is not confined to contexts where veiling is prevalent. That said, extrapolation to Europe must be made with caution: veiling rates in Turkey remain much higher than in most European countries, and veiled women in Europe may face different challenges, including labor market discrimination and social exclusion, beyond the formal institutional barriers we study here. Thus, our evidence is best viewed as suggestive for Muslim-minority contexts.

Moreover, our results indicate a substitution effect of the reform: veiled women entered public employment while non-veiled women shifted toward private sector jobs, implying a reallocation of employment opportunities rather than the creation of new ones. This raises a central question for interpreting the reform's implications: were the veiled women who gained access to public employment after the repeal at least as qualified as those they replaced, or did the reform simply facilitate politically motivated hiring?

While our earlier evidence shows that the increase in public employment among veiled women is not concentrated in politically aligned regions, suggesting limited local patronage, this pattern does not rule out favoritism operating at the national level, for instance through uniform recruitment practices that advantaged veiled applicants across all regions. To examine this possibility, we restrict the sample to women employed in the public sector and estimate Equation (1) for three education–occupation match indicators based on ISCO skill levels: (i) overqualified (education exceeds the occupation's skill requirement), (ii) underqualified (education below the requirement), and (iii) mismatch (either condition). These indicators serve as proxies for allocative efficiency in the public sector. If political favoritism had displaced more qualified candidates, we would expect an increase in underqualification and overall mismatch among veiled women after the reform, relative to comparable non-veiled women. The coefficients on $Veiled \times Post$ reported in Online Appendix Table A11, captures the differential before–after change in match quality between veiled and non-veiled public employees. In Panel A, we find no significant change in overqualification or overall mismatch, but a statistically significant 1.9-percentage-point increase in underqualification. The magnitude is small and does not translate into higher overall mismatch. In Panel B, which adjusts for group-specific trends, the underqualification effect disappears, whereas overqualification and mismatch both decline significantly. On one hand, the fall in mismatch suggests a closer alignment between workers' education levels and job requirements, consistent with improved allocative matching. On the other hand, the reduction in overqualification could also indicate that the average education level of veiled entrants was lower relative to the jobs they obtained, implying a decline in the stock of human capital within these positions.

[24]In countries such as Egypt, Tunisia, Azerbaijan, Kazakhstan, Kosovo, and Kyrgyzstan, various forms of veiling bans persist in schools or public spaces.

Interpreting these indicators jointly, the results point to a compositional shift toward a more balanced, but not necessarily more skilled public sector workforce. Thus, while the reform does not appear to have generated large efficiency losses through political favoritism or severe misallocation, it may have slightly reduced average qualification levels as access broadened. In column 4, we also find that veiled women earn slightly less than comparable non-veiled public employees after the reform. Such a pattern is inconsistent with a favoritism channel: if politically motivated hiring were dominant, we would expect higher wages among the favored group. Instead, the modest negative earnings gap aligns with the seniority-based wage structure of the Turkish public sector, where pay progression depends primarily on tenure. Because most veiled women entered only after the reform, their shorter tenure mechanically implies lower earnings, not lower productivity.

Taken together, these results suggest that the repeal broadened access to public employment without evidence of politically driven favoritism or major short-run efficiency losses. The decline in overqualification may reflect a normalization of the workforce composition rather than a deterioration in job–skill match quality. Future work linking changes in public sector employment composition to sectoral outcomes, such as student achievement or patient health indicators, could provide a more direct assessment of the reform's longer-term implications for public sector productivity.

Beyond the labor market, a natural question is whether the repeal of the headscarf ban also affected family outcomes. In Online Appendix Table A12, we examine marriage and divorce probabilities before and after the repeal of headscarf ban. Panel A shows no effects, but when we adjust for group-specific trends in Panel B, the estimates suggest a decline in marriage and a small increase in divorce among veiled women. Robustness checks following Rambachan and Roth (2023) support this pattern (see Online Appendix Figure A2), but the divergence between Panels A and B indicates that the parallel-trends assumption may be violated for marriage outcomes, making these results more suggestive than definitive. Even so, the pattern is intriguing: as access to public sector jobs expands, veiled women may delay or forgo marriage due to rising opportunity costs, while the increase in divorce could reflect stronger bargaining power within the household. These findings point to potential empowerment effects of lifting institutional restrictions within the family, extending beyond women's labor market outcomes. Examining these channels with richer longitudinal data on marriage timing, household decision-making, and veiling behavior remains a promising direction for future research.

A further avenue for research concerns the dynamics of veiling itself. Our empirical design treats veiling as time-invariant, but the broader question of how labor market opportunities affect women's veiling decisions remains largely unexplored. Individual-level panel data from sources such as KONDA could allow researchers to study whether the repeal of the ban altered women's veiling choices directly, and more generally, how changes in labor market opportunities interact with religious expression.[25]

6 Conclusion

This paper examines how the repeal of the headscarf ban in public institutions in Turkey affects women's labor market outcomes. Using two complementary datasets, the Turkish Household Labor Force Survey and the Turkish Demographic and Health Survey, we predict

[25] We attempted a preliminary analysis using the 2008 and 2013 DHS, but the lack of variation in women's employment probabilities across these two points prevented us from drawing meaningful conclusions. More frequent survey waves capturing a longer time period would be needed to examine this question systematically.

and impute women's veiling status from observable characteristics through statistical matching and machine learning methods. We then compare the labor market outcomes of veiled and non-veiled women before and after the 2013 reform within a difference-in-differences framework.

The results reveal three main patterns. First, the repeal substantially increases public sector employment among women predicted to be veiled. This increase reflects both a higher probability of employment and a shift away from self-employment and unpaid family work, suggesting that the reform reshaped women's occupational choices rather than simply expanding jobs. Second, most of the gains for veiled women occur in professional and associate-professional roles, such as teachers, nurses, and administrative staff, which typically require higher levels of education. Consistent with this pattern, the effects are much stronger among more-educated veiled women. Third, non-veiled women experience a decline in public sector employment and a corresponding rise in private sector work, indicating a substitution of job opportunities from non-veiled to veiled women rather than an overall increase in public hiring. Analysis of education–occupation matching shows no systematic signs of notable efficiency losses in the public sector.

The heterogeneity analysis across regions that differ in veiling prevalence and political alignment provides deeper insight into the reform's impacts. Veiled women living in low-veiling regions, who are on average less religious, respond more strongly to the reform. This pattern indicates that the observed policy effects are not driven by religiosity, and that the estimated impact is likely a lower bound on the true causal effect of lifting the ban. Veiled women in regions under stronger Islamist political influence also experience larger gains, particularly in overall and public sector employment. However, the effect on public sector employment remains positive and statistically significant even in less Islamist regions. These results suggest that while Islamist-leaning municipalities may have facilitated more access for veiled women, the main driver of the observed gains is the nationwide removal of the institutional barrier that had previously excluded them from public employment.

Future research can build on this study in several directions. One avenue is to examine more closely the reform's implications for public sector efficiency and output quality by collecting detailed data to measure productivity at the institutional or sectoral level. Another promising direction is to investigate marriage market outcomes in greater depth. The current analysis provides suggestive evidence that the reform increased women's empowerment, with veiled women delaying marriage and experiencing a modest rise in divorce. Access to richer longitudinal data would allow more precise estimates of effects on marriage timing and intra-household decision-making. Finally, future work could explore how the reform influenced the educational choices of younger religious women and whether subsequent changes in educational attainment feed back into veiling practices over time. Finally, future work could explore whether and how the repeal affects the education choices of young religious women in the long run.

Taken together, these findings highlight that easing institutional restrictions on religious expression can significantly improve women's economic inclusion. This suggests that policy interventions expanding institutional access, rather than attempts to reshape individual identities or cultural norms, may be more effective in improving economic opportunities for underrepresented groups. While the prediction of veiling status and the relatively short post-reform period impose some limitations, the results are robust across specifications and offer credible short-run evidence.

References

Abdelgadir, A. and V. Fouka (2020). Political secularism and Muslim integration in the West: Assessing the effects of the French headscarf ban. *American Political Science Review 114*(3), 707–723.

Abdelhadi, E. (2019). The Hijab and Muslim women's employment in the United States. *Research in Social Stratification and Mobility 61*, 26–37.

Abou Daher, C., E. Field, K. Swanson, and K. Vyborny (2025). Drivers of change: Employment responses to the lifting of the Saudi female driving ban. *American Economic Review 115*(9), 3248–3271.

Aha, D. W., D. Kibler, and M. K. Albert (1991). Instance-based learning algorithms. *Machine learning 6*, 37–66.

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *The Quarterly Journal of Economics 115*(3), 715–753.

Akoglu, K. S. (2015). Piecemeal freedom: Why the headscarf ban remains in place in Turkey. *Boston College International and Comparative Law Review 38*, 277.

Aksoy, O. and D. Gambetta (2016). Behind the veil: The strategic use of religious garb. *European Sociological Review 32*(6), 792–806.

Aksoy, O. and D. Gambetta (2021). The politics behind the veil. *European Sociological Review 37*(1), 67–88.

Alesina, A. and P. Giuliano (2015). Culture and institutions. *Journal of economic literature 53*(4), 898–944.

Alesina, A., P. Giuliano, and N. Nunn (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics 128*(2), 469–530.

Alvi, A. (2013). Concealment and revealment: The Muslim veil in context. *Current Anthropology 54*(2), 177–199.

Anıl, E. (2002). *The new legal status of women in Turkey.* Women for Women's Human Rights-New Ways.

Athey, S. and G. W. Imbens (2019). Machine learning methods that economists should know about. *Annual Review of Economics 11*(1), 685–725.

BBC News (2010). Turkey drops Islamic headscarf ban at universities. Accessed: 2025-01-06.

BBC News (2017). Turkey's Erdogan backs removing headscarf ban in military. Accessed: 2025-01-03.

Bianet (2010). Headscarf ban at universities softened. Accessed: 2025-01-06.

Breiman, L. (1996). Bagging predictors. *Machine learning 24*, 123–140.

Breiman, L. (2001). Random forests. *Machine learning 45*, 5–32.

Bursztyn, L., A. L. González, and D. Yanagizawa-Drott (2020). Misperceived social norms: Women working outside the home in Saudi Arabia. *American Economic Review 110*(10), 2997–3029.

Carvalho, J.-P. (2013). Veiling. *The Quarterly Journal of Economics 128*(1), 337–370.

Cattaneo, M. D., N. Idrobo, and R. Titiunik (2019). *A practical introduction to regression discontinuity designs: Foundations*. Cambridge University Press.

Chen, A. Y. and J. McCoy (2024). Missing values handling for machine learning portfolios. *Journal of Financial Economics 155*, 103815.

Cindoğlu, D. (2011). *Headscarf ban and discrimination: Professional headscarved women in the labor market*. Tesev Publications İstanbul, Turkey.

Cindoglu, D. and G. Zencirci (2008). The headscarf in Turkey in the public and state spheres. *Middle Eastern Studies 44*(5), 791–806.

Corekcioglu, G. (2021). Unveiling the effects of a headscarf ban: Evidence from municipal jobs in Turkey. *Journal of Comparative Economics 49*(2), 382–404.

Cortes, C. (1995). Support-vector networks. *Machine Learning*.

Dhar, D., T. Jain, and S. Jayachandran (2022). Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India. *American Economic Review 112*(3), 899–927.

Dildar, Y. (2015). Patriarchal norms, religion, and female labor supply: Evidence from Turkey. *World Development 76*, 40–61.

Drouhot, L., P. Simon, and V. Tiberj (2023). La diversité religieuse en france: transmissions intergénérationnelles et pratiques selon les origines. *Références-INSEE*, 39–48.

Drouhot, L. G. (2021). Cracks in the melting pot? Religiosity and assimilation among the diverse Muslim population in France. *American Journal of Sociology 126*(4), 795–851.

Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic literature 50*(4), 1051–1079.

Erten, B. and P. Keskin (2018). For better or for worse? Education and the prevalence of domestic violence in Turkey. *American Economic Journal: Applied Economics 10*(1), 64–105.

Euronews (2017). Turkey purge continues with 7000 public sector workers dismissed. Accessed: 2025-01-03.

Fernández, R. and A. Fogli (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American economic journal: Macroeconomics 1*(1), 146–177.

Fernández-Reino, M., V. Di Stasio, and S. Veit (2023). Discrimination unveiled: A field experiment on the barriers faced by Muslim women in Germany, the Netherlands, and Spain. *European Sociological Review 39*(3), 479–497.

Field, E., S. Jayachandran, and R. Pande (2010). Do traditional institutions constrain female entrepreneurship? A field experiment on business training in India. *American Economic Review 100*(2), 125–129.

Field, E., R. Pande, N. Rigol, S. Schaner, and C. Troyer Moore (2021). On her own account: How strengthening women's financial control impacts labor supply and gender norms. *American Economic Review 111*(7), 2342–2375.

Fleischmann, F. and K. Phalet (2012). Integration and religiosity among the Turkish second generation in Europe: A comparative analysis across four capital cities. *Ethnic and Racial Studies 35*(2), 320–341.

Fouka, V. (2020). Backlash: The unintended effects of language prohibition in US schools after World War I. *The Review of Economic Studies 87*(1), 204–239.

Gazeaud, J., N. Khan, E. Mvukiyehe, and O. Sterck (2023). With or without him? Experimental evidence on cash grants and gender-sensitive trainings in Tunisia. *Journal of Development Economics 165*, 103169.

Ghumman, S. and L. Jackson (2010). The downside of religious attire: The Muslim headscarf and expectations of obtaining employment. *Journal of organizational behavior 31*(1), 4–23.

Gogas, P. and T. Papadimitriou (2021). Machine learning in Economics and Finance. *Computational Economics 57*, 1–4.

Goldstein, I., C. S. Spatt, and M. Ye (2021). Big data in Finance. *The Review of Financial Studies 34*(7), 3213–3225.

Göle, N. (1996). *The forbidden modern: Civilization and veiling.* University of Michigan Press.

Gonzales, M. C., M. S. Jain-Chandra, M. K. Kochhar, and M. M. Newiak (2015). *Fair play: More equal laws boost female labor force participation.* International Monetary Fund.

Güneş, P. M. (2016). The impact of female education on teenage fertility: Evidence from Turkey. *The BE Journal of Economic Analysis & Policy 16*(1), 259–288.

Guveli, A. (2011). Social and economic impact of the headscarf ban on women in Turkey. *European Societies 13*(2), 171–189.

Hallward-Driemeier, M. and O. Gajigo (2015). Strengthening economic rights and women's occupational choice: The impact of reforming Ethiopia's family law. *World Development 70*, 260–273.

Heller, S. B., B. Jakubowski, Z. Jelveh, and M. Kapustin (2024). Machine learning can predict shooting victimization well enough to help prevent it. *Review of Economics and Statistics*, 1–45.

Höjelid, S. (2010). Headscarves, judicial activism, and democracy: The 2007–8 constitutional crisis in Turkey. *The European Legacy 15*(4), 467–482.

Hyland, M., S. Djankov, and P. K. Goldberg (2020). Gendered laws and women in the workforce. *American Economic Review: Insights 2*(4), 475–490.

Jacquet, A. and S. Montpetit (2023). Veiling and the Economic Integration of Muslim women in France.

Jayachandran, S. (2021). Social norms as a barrier to women's employment in developing countries. *IMF Economic Review 69*(3), 576–595.

Joslin, K.-E. and F. M. Nordvik (2021). Does religion curtail women during booms? Evidence from resource discoveries. *Journal of Economic Behavior & Organization 187*, 205–224.

Karahan, H. and N. Tuğsuz (2022). Political climate and the headscarf issue in Turkey. *Insight Turkey 24*(1), 211–228.

Kırdar, M. G., M. Dayıoğlu, and İ. Koç (2018). The effects of compulsory-schooling laws on teenage marriage and births in Turkey. *Journal of Human Capital 12*(4), 640–668.

KONDA (2019). July 2019 Barometer: Public opinion research report. Accessed: 2025-01-06.

Maurin, E. and N. Navarrete H (2023). Behind the veil: The effect of banning the Islamic veil in schools. *Economic Policy 38*(113), 63–98.

Merlino, L. P. and E. Yurdakul (2024). Women's education and household labor supply: Causal evidence from Turkey.

Mullainathan, S. and Z. Obermeyer (2022). Diagnosing physician error: A machine learning approach to low-value health care. *The Quarterly Journal of Economics 137*(2), 679–727.

Mullainathan, S. and J. Spiess (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives 31*(2), 87–106.

OECD (2025). Government at a glance 2025.

Oh, S. (2023). Does identity affect labor supply? *The American Economic Review 113*(8), 2055–2083.

Rambachan, A. and J. Roth (2023). A more credible approach to parallel trends. *Review of Economic Studies 90*(5), 2555–2591.

Ridder, G. and R. Moffitt (2007). The econometrics of data combination. *Handbook of econometrics 6*, 5469–5547.

Rosenbaum, P. R. and D. B. Rubin (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika 70*(1), 41–55.

Roy, S. (2019). Discriminatory laws against women: A survey of the literature. *World Bank Policy Research Working Paper* (8719).

Rubin, D. B. (1986). Statistical matching using file concatenation with adjusted weights and multiple imputations. *Journal of Business & Economic Statistics 4*(1), 87–94.

Rumelhart, D. E., G. E. Hinton, and R. J. Williams (1986). Learning representations by back-propagating errors. *Nature 323*(6088), 533–536.

Saleh, M. and J. Tirole (2021). Taxing identity: Theory and evidence from early Islam. *Econometrica 89*(4), 1881–1919.

Shofia, N. (2022). Why veil? Religious headscarves and the public role of women. *Religious Headscarves and The Public Role of Women (July 21, 2022)*.

Uğur, Z. B. (2020). Unveiled: The effect of the headscarf ban on women's tertiary education in Turkey. *Feminist Economics 26*(2), 187–217.

Wiltse, E. Ç. (2008). The gordian knot of Turkish politics: Regulating headscarf use in public. *South European Society and Politics 13*(2), 195–215.

A Online Appendix

A.1 Effects of the 1997 Education Reform and The 2022 Civil Code on Veiling Decision

To identify the causal effect of the 1997 education reform, we assume that aside from the increase in education induced by the 1997 reform, women born just before and just after January 1987 are otherwise comparable in observable and unobservable characteristics. Based on this assumption, we exploit the exogenous variation in the birth months and years generated by the change in the compulsory schooling law in Turkey. Our empirical strategy is based on a regression discontinuity (RD) design, comparing the probability of being veiled of women born in or after January 1987 (who must complete eight years of education) to those women was born before January 1987 (who could drop out after five years of compulsory education). Formally, we estimate the following equation:

$$y_i = \alpha_0 + \alpha_1 D_i + \alpha_2 X_i + \alpha_3 (D_i \times X_i) + \gamma' Z_i + \varepsilon_i, \tag{2}$$

where y_i denotes the probability of veiling for woman i, D_i is a binary indicator equal to one if the woman was born in or after January 1, 1987. The term X_i is the normalized running variable around January 1987, defined as the number of months from the cutoff date, flexibly capturing smooth trends in outcomes as a function of birth timing.[26] The interaction term $X_i \times D_i$ allows the slope of these trends to differ on either side of the cutoff. Our coefficient of interest, α_1, captures the average discontinuity at the cutoff and is interpreted as the causal effect of exposure to the 1997 reform for women born just after January 1987, relative to those born just before.

Estimation of model (2) involves specifying the functional form of the running variable, X_i. We model this function as a linear polynomial, determining an optimal bandwidth using the algorithm proposed by Cattaneo et al. (2019). To address potential fuzziness in the treatment status, we also estimate Eq. (2) by excluding women born within three months before or after January 1, 1987. The rationale for using a ± 3 month window is that children born closest to the cutoff date are the most likely, if any, to have started school a year earlier or later than their assigned cohort, due to possible deviations from standard enrollment age.[27] By omitting births within this window, we ensure cleaner separation between cohorts exposed to the different schooling regimes with minimal sample loss.

We control for a vector of variables Z_i comprising dummies for childhood place of residence, mother tongue, whether the woman's mother is literate, dummies for father's education level, the woman's month-of-birth fixed effects, and region-of-residence fixed effects. Standard errors are clustered at the month-year of birth level.

To identify the causal effect of the 2002 Civil Code reform, we assume that aside from the restrictions induced by the increase in the minimum legal marriage age, women born just before and just after January 1985 are otherwise comparable in observable and unobservable characteristics. Based on this assumption, we exploit exogenous variation in birth timing around the January 1985 cutoff, which determines the duration of exposure to the new Civil Code in Turkey.

[26] The running variable is centered at the January 1987 cutoff; positive values indicate birth months after the cutoff, and negative values indicate months before.

[27] For example, a child born in February 1987 would typically start school in September 1993, the year they turn six. However, due to enrollment practices at the time, such a child could have started a year early—in September 1992, at age five and a half. In contrast, a child born in August 1986 would be too old to delay entry by a full year, making early or late enrollment less likely. This justifies our focus on the $\pm$3-month window, where deviations from standard enrollment age are more plausible.

Our empirical strategy is based on a regression discontinuity (RD) design with dose intensity, comparing veiling probabilities of women born between 1983 and 1986. Women born before January 1985 were not affected by the reform, while those born between January 1985 and December 1986 were affected to varying degrees: a woman born in January 1985 was exposed for one month, whereas a woman born in December 1986 was exposed for 24 months.

$$y_i = \beta_0 + \beta_1 \, \text{ExposureMonths}_i + \beta_2 X_i + \gamma' Z_i + \varepsilon_i \tag{3}$$

where y_i denotes the veiling probability for woman i, ExposureMonths_i is the number of months for which she was legally restricted by the reform (equal to zero for those born before January 1985, and between 1 and 24 for those born between January 1985 and December 1986), and X_i is the normalized running variable defined as the number of months from the January 1985 cutoff.[28] Our coefficient of interest, β_1, captures the causal effect of an additional month of exposure to the higher marriage age requirement on the probability of veiling. In practice, we report this effect both per month and scaled to a 12-month increase in exposure. We restrict the sample around a $\pm$24-month window around the cutoff, and also report robustness checks with alternative cohorts.

A.2 Machine learning techniques

Machine learning (ML) techniques are increasingly applied to address missing data problems (Chen and McCoy, 2024) and to improve out-of-sample predictions (Athey and Imbens, 2019). Compared with other statistical methods, ML approaches offer advantages such as flexible functional forms, greater computational efficiency and higher accuracy (Mullainathan and Spiess, 2017). Given the common challenges of poor data quality in various fields, a variety of ML methods have been proposed to tackle missing value issues in different contexts (Gogas and Papadimitriou, 2021). Applications in economics include healthcare data (Mullainathan and Obermeyer, 2022), asset pricing (Goldstein et al., 2021), and other areas.

Machine learning approaches are typically categorized into supervised, unsupervised, and semi-supervised learning. The key distinction lies in whether labeled datasets are available—that is, whether clear relationships between input features and corresponding output are established. In this paper, we use supervised learning, as we have observed demographic characteristics as inputs and veiling status as outputs derived from the DHS dataset. Supervised learning achieves high accuracy when the training dataset is abundant and well-labeled. Since the DHS is a high-quality, nationally representative, and randomized dataset, we can reasonably expect strong performance from the applied ML techniques.

The ML procedure is divided into two main steps: training and prediction. In the training phase, we analyze the DHS dataset to identify patterns between observed demographic characteristics and the headscarf-wearing behavior of female respondents. Using a selected ML algorithm, we develop a model capable of predicting headscarf-wearing outcomes for unseen inputs. To train the model, we split the DHS dataset into training and test sets. The algorithm uses the training set to fit the model, and its predictions are compared against the true outputs in the test set. The model's parameters are adjusted iteratively to minimize the difference between its predictions and the actual outcomes. In the prediction phase, the optimized model is applied to the HLFS dataset to predict unknown veiling outcomes based on observed demographic inputs. Several key algorithms have demonstrated strong performance in recent studies.

[28] The running variable is centered at January 1985; positive values indicate birth months after the cutoff, and negative values indicate months before.

Ensemble learning combines multiple models to improve accuracy, robustness, and flexibility by aggregating their predictions (Breiman, 1996). The approach is particularly effective when individual algorithms capture different data patterns or when a single model lacks sufficient accuracy. Besides, it helps mitigate overfitting in small or noisy datasets. A variety of models can be incorporated into ensemble learning, with Support Vector Machines, Random Forests, K-Nearest Neighbors, and Neural Networks among the most common.

For our robustness checks in scarf prediction, we first implement ensemble learning by combining Support Vector Machines, Random Forests, and K-Nearest Neighbors. We then apply each of these models separately, and finally include a Neural Network model. The following sections provide a detailed discussion of these models.

A.2.1 Support Vector Machines

Support Vector Machines (SVM) is a powerful algorithm for data classification (Cortes, 1995). The core principle is to find an optimal decision boundary, known as a hyperplane, that separates two classes of data points. The objective is to maximize the margin—the distance between the hyperplane and the nearest data points from each class, called support vectors—while ensuring accurate classification.[29]

Consider a two-class classification problem with training data $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$, where $\mathbf{x}_i$ are feature vectors and $y_i \in \{-1, 1\}$ are class label.

In the primal form of SVM, the goal is to find a hyperplane that optimally separates the two classes. The decision function is a linear function defined as:

$$f(x) = \mathbf{w} \cdot \mathbf{x} + b$$

where $\mathbf{w} \in \mathbb{R}^n$ is the weight vector (orthogonal to the decision hyperplane), $\mathbf{x} \in \mathbb{R}^n$ is the input feature vector, and $b \in \mathbb{R}$ is the bias term. The dot product $\mathbf{w} \cdot \mathbf{x}$ represents the projection of the data point x onto vector w, which plays a key role in defining the decision boundary.

For classification, the predicted class is determined by the sign of $f(x)$. If $f(x) > 0$, predict class is $+1$; and if $f(x) < 0$, predict class is -1.

The primal optimization problem aims to find the weight vector w and bias term b that define the optimal separating hyperplane. This is formulated as:

$$\min_{\mathbf{w},b} \frac{1}{2}\|\mathbf{w}\|^2$$

subject to the classification constraints:

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 \quad \text{for all } i$$

The objective function $\frac{1}{2}\|\mathbf{w}\|^2$ is minimized to maximize the margin between the two classes, while the constraints ensures that each data point is classified correctly.

When the data is non-linearly separable, the approach is to map the data into a higher-dimensional space using a feature mapping function $\phi(x)$. In this higher-dimensional space, the data becomes linearly separable, and the decision function is given by:

$$f(x) = \mathbf{w} \cdot \phi(\mathbf{x}) + b$$

[29] Maximizing the margin helps create a more robust decision boundary, making it less sensitive to noise and outliers.

Instead of explicitly computing $\phi(x)$, the kernel trick is used to replace the dot product $\mathbf{w} \cdot \mathbf{x}_i$ with a kernel function $K(\mathbf{x}_i, \mathbf{x}_j)$, which directly computes the similarity between data points in the higher-dimensional space.

The dual form of the SVM optimization problem is then written in terms of the Lagrange multipliers α_i:

$$\max_{\alpha_i} \left(\sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \right)$$

subject to the constraints:

$$\sum_{i=1}^{N} \alpha_i y_i = 0 \quad \text{and} \quad 0 \leq \alpha_i \leq C$$

Here the C is a regularization parameter that controls the trade-off between maximizing the margin and minimizing classification errors. $K(\mathbf{x}_i, \mathbf{x}_j)$ is the kernel function. We have several commonly used kernel functions include: linear, polynomial and radial basis function Kernel.

As shown, a key advantage of SVM is its flexibility in handling complex data structures by employing different kernel functions, making it well-suited for nonlinear classification problems. After performing a grid search to optimize model parameters for maximum training accuracy, we configure the SVM model as follows:

```
svm_model = SVC(kernel='rbf', C=10, gamma='scale')
```

Here, we use the Radial Basis Function (RBF) kernel to capture nonlinear decision boundaries. The gamma parameter is set to 'scale', adapting to the data to balance complexity and prevent overfitting or underfitting. Additionally, the regularization parameter $C = 10$, emphasizing classification accuracy over margin maximization, allowing the model to handle misclassified points with lower tolerance.

A.2.2 Random Forest

Random Forest (RF) is a powerful machine learning algorithm known for its ability to identify important features and its robustness to noise and outliers, making it highly resistant to overfitting (Breiman, 2001). It is a reliable choice for a variety of tasks, especially classification. The implementation involves several key steps:

The first step is "Bootstrap Sampling". We generate multiple bootstrap samples from the original dataset. Each sample $D_{\text{bootstrap}}$ is a random subset created by sampling with replacement. [30]

$$D_{\text{bootstrap}} = \{\text{Sample from } D \text{ with replacement}\}$$

where D is the original dataset.

In the second step "Training Decision Trees", a decision tree is independently trained on each bootstrap sample $D_{\text{bootstrap}}$. Here's how the training process works:

At each node of the tree, the algorithm evaluates which feature and corresponding split threshold best separates the data into subsets. This is done by assessing potential splits using impurity measures such as Gini impurity or entropy. The goal is to find the feature and threshold that minimize the impurity, creating more homogeneous subsets.

$$Gini(D) = 1 - \sum_{i=1}^{C} p_i^2$$

[30] Some data points may be repeated, and some may be left out.

where p_i is the proportion of class i in the node, and C is the number of classes. A lower Gini impurity means the node is more pure (i.e., most of the samples in the node belong to the same class), while a higher Gini impurity indicates that the samples are more mixed across different classes.

Once the best feature and threshold are identified for the current node, the dataset is split into two child nodes based on the chosen feature and value. This process is repeated recursively for each child node in the same way, each time selecting the best feature to further split the data, until a stopping criterion is met. Some common stopping criteria include: maximum depth of the tree, minimum samples per leaf, minimum impurity split and etc.

The third step is "Prediction". Once all trees are trained, predictions are made by aggregating the outputs of all trees. A "majority vote" is taken to determine the final class label. The class that receives the most votes from the individual trees is selected as the final prediction:

$$\hat{y} = \text{mode}(\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_T)$$

where $\hat{y}_i$ is the predicted class from tree i, and T is the total number of trees.

After performing a grid search to find the optimal parameters for the best accuracy, we configure the RF model as follows:

- `n_estimators` = 100: This specifies the number of trees in the random forest. A higher number improves the model's robustness but also increases computation time.

The following three parameters are stopping criteria that control how the individual decision trees are trained:

- `max_depth` = 10: This sets the maximum depth of each tree. Limiting depth prevents the trees from becoming too complex and overfitting to the training data.

- `min_samples_split` = 5: This defines the minimum number of samples required to split an internal node. Nodes with fewer than 5 samples cannot be split further. This helps in reducing overfitting by ensuring that splits are based on sufficiently large subsets of data.

- `min_samples_leaf` = 4: This controls the minimum number of samples required in each leaf node after a split. It ensures each leaf contains at least 4 samples, which prevents overly specific splits and helps avoid overfitting.

```
rf_model = RandomForestClassifier(
    n_estimators=100,
    max_depth=10,
    min_samples_leaf=4,
    min_samples_split=5
)
```

A.2.3 K-Nearest Neighbors and Neural Networks

As part of our sensitivity analysis, we implement two other commonly used ML models and present the diff-in-diff results using imputed data from these models.

K-Nearest Neighbor (K-NN) is a simple yet powerful instance-based algorithm (Aha et al., 1991). Unlike other models, K-NN doesn't require an explicit training phase. Instead, it stores the training dataset and uses it directly for prediction. To predict the output for a new data point, K-NN calculates the distance between the new point and all existing points in the training set, typically using Euclidean distance or another distance measure. It then identifies the K-nearest neighbors and classifies or regresses based on the labels or values of these neighbors. We select the optimal parameters for the model as follows:

```
knn_model = KNeighborsClassifier(
    n_neighbors=15,
    algorithm = 'auto',
    leaf_size=30,
    metric = 'minkowski',
    n_jobs=-1,
    p=1,
    weights = 'uniform'
)
```

Neural Networks (NN) are powerful machine learning models inspired by the structure and functioning of the human brain. They consist of interconnected nodes, or neurons, organized into layers, which allow the model to recognize patterns and learn complex relationships in data. Neural networks are designed to learn from data through a series of connected layers, with each layer performing mathematical operations to transform input data into useful outputs (Rumelhart et al., 1986).

For our sensitivity analysis, we establish a neural network model with the following architecture:

- `Input Layer`: Contains $X\_train.shape[1]$ neurons, equal to the number of input features.

- `Hidden Layers`: Two hidden layers with 64 and 32 neurons. These layers use the ReLU activation functions to introduce non-linearity.

- `Output Layer`: one neuron with a sigmoid activation function to produce binary classification probabilities.

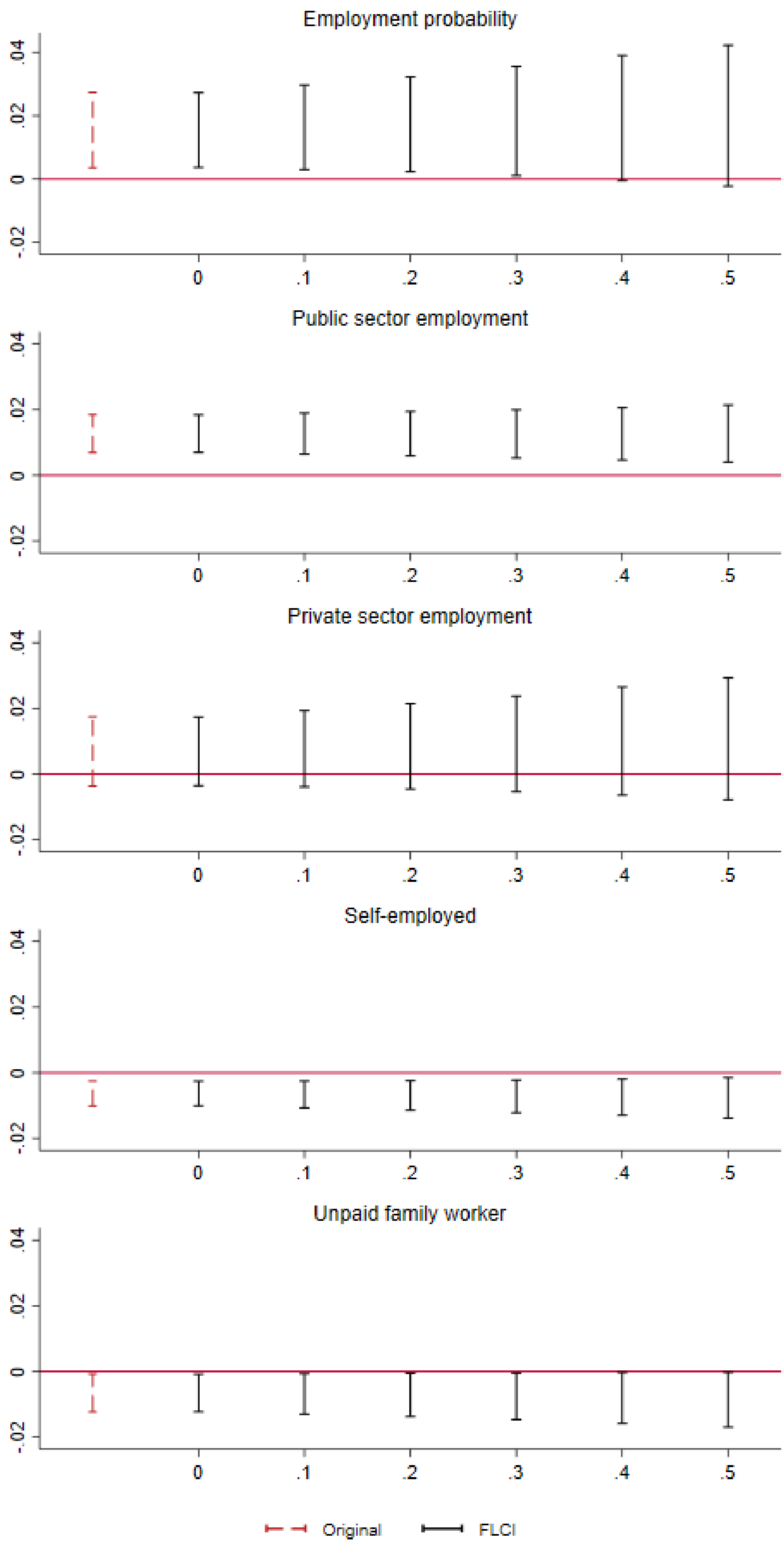
The model is compiled using the Adam optimizer and binary cross-entropy loss, and we monitor accuracy during training to assess performance.

```
model = Sequential()
model.add(Dense(
    64, input_dim=X_train.shape[1], activation='relu'))
model.add(
    Dense(32, activation='relu'))
model.add(
    Dense(1, activation='sigmoid'))

model.compile(
    optimizer='adam',
    loss='binary_crossentropy',
    metrics=['accuracy'])
```

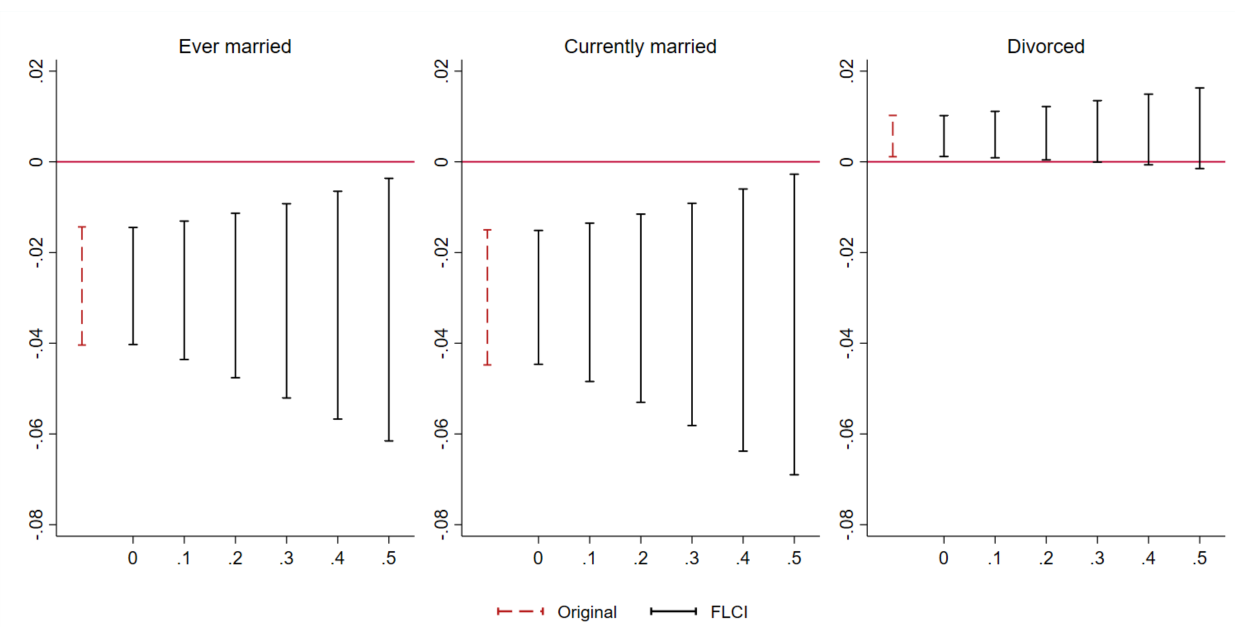
A.3 Figures

Figure A1: Effects on labor market outcomes: sensitivity analysis



Notes: Figures display sensitivity analysis of estimated treatment effects on labor market outcomes of women to potential violations of the parallel trends assumptions outlined in Rambachan and Roth (2023). The red dashed bar in each panel represents the 95% confidence interval of our baseline estimates. The black solid bars represent corresponding 95% confidence intervals when allowing for pre-reform period violations of parallel trends of up to M, indicating the largest allowable change in the slope of an underlying linear trend between two consecutive years.

Figure A2: Effects on marriage probability: sensitivity analysis



Notes: Figures display sensitivity analysis of estimated treatment effects on the marriage probability of women to potential violations of the parallel trends assumptions outlined in Rambachan and Roth (2023). The red dashed bar in each panel represents the 95% confidence interval of our baseline estimates. The black solid bars represent corresponding 95% confidence intervals when allowing for pre-reform period violations of parallel trends of up to M, indicating the largest allowable change in the slope of an underlying linear trend between two consecutive years.

A.4 Tables

Table A1: *Occupations in the public sector (%)*

	Men	Women
Legislators and Senior Officials	4.06	3.02
Professionals	41.09	57.81
Technicians and Associate Professionals	12.49	12.71
Clerks	14.69	17.34
Service Workers and Shop and Market Sales Workers	15.14	6.03
Skilled Agricultural and Fishery Workers	0.29	0.11
Craft and Related Trades Workers	2.59	0.37
Plant and Machine Operators and Assemblers	2.44	0.08
Elementary Occupations	7.19	2.54
Total	100.00	100.00

Notes: The share of men and women in each occupation is calculated using employed samples from the Turkish Household Labor Force Survey. We follow the SIC-92 classifications.

Table A2: *Summary statistics using the HLFS dataset*

	All sample	Non-veiled	Veiled
	Mean (S.D) (1)	Mean (S.D) (2)	Mean (S.D) (3)
Panel A: Background characteristics			
Age	32.88	31.46	33.79
	(8.906)	(8.564)	(9.000)
Native	0.978	0.960	0.989
	(0.148)	(0.196)	(0.105)
University degree	0.151	0.322	0.042
	(0.358)	(0.467)	(0.200)
High school degree and above	0.352	0.640	0.169
	(0.478)	(0.480)	(0.374)
Primary school degree and below	0.494	0.228	0.664
	(0.499)	(0.420)	(0.472)
Married	0.713	0.582	0.797
	(0.452)	(0.493)	(0.402)
Household size	3.925	3.235	4.271
	(1.914)	(1.310)	(2.070)
Panel B: Labor market outcomes			
Employed	0.330	0.395	0.289
	(0.470)	(0.489)	(0.453)
Employed as wage earner	0.222	0.339	0.147
	(0.415)	(0.474)	(0.354)
Public sector employment	0.051	0.102	0.019
	(0.221)	(0.302)	(0.138)
Self-employed	0.034	0.0325	0.036
	(0.184)	(0.177)	(0.187)
Unpaid family worker	0.083	0.044	0.108
	(0.276)	(0.204)	(0.311)
Obs.	933,810	328,862	604,948

Notes: Data are from the Turkish HLFS 2010-17. The sample covers all women aged 18-49. The table presents the means, and standard deviations (in parenthesis) of selected variables for the sample of all women in column 1, for the sample of women who are predicted to be non-veiled women in column 2, and for those predicted to be veiled in column 3, respectively. Data are weighted using the cross-sectional weights for the wave at which the outcome was measured.

Table A3: *Summary statistics for veiled and non-veiled women from the 2013 DHS*

	Non-veiled	Veiled	
	(1) Mean (S.D)	(2) Mean (S.D)	(3) Difference (1)-(2)
Panel A: Background characteristics			
Age	31.17	33.61	-2.446***
	(8.631)	(8.799)	(0.234)
Native	0.951	0.989	-0.037***
	(0.215)	(0.106)	(0.005)
University degree	0.387	0.067	0.320***
	(0.487)	(0.250)	(0.011)
High school degree and above	0.468	0.109	0.359***
	(0.499)	(0.311)	(0.012)
Primary school degree and below	0.204	0.645	-0.441***
	(0.008)	(0.007)	(0.011)
Married	0.616	0.841	-0.225***
	(0.487)	(0.366)	(0.012)
Household size	3.752	5.067	-1.314***
	(1.335)	(2.296)	(0.047)
Panel B: Labor market outcomes			
Employed	0.427	0.268	0.159***
	(0.495)	(0.443)	(0.013)
Employed as wage earner	0.334	0.0839	0.250***
	(0.472)	(0.277)	(0.011)
Employed in public sector	0.095	0.012	0.083***
	(0.293)	(0.108)	(0.006)
Self-employed	0.038	0.054	-0.016***
	(0.190)	(0.227)	(0.005)
Unpaid family worker	0.022	0.091	-0.068***
	(0.147)	(0.287)	(0.005)
Obs.	2,793	6,005	

Notes: The data are from the 2013 Turkish Demographic and Health Survey. Sample covers all women aged 18-49. The table presents the means, and standard deviations (in parenthesis) of selected variables for non-veiled women in column 1, and for veiled women in column 2, and the difference between non-veiled and veiled women in column 3. ***, significant at the 1 percent level; **, significant at the 5 percent level; *, significant at the 10 percent level.

Table A4: *Effects on labor market outcomes of non-veiled women*

Outcome	Employment probability (1)	Public sector employment (2)	Private sector employment (3)	Self-employed (4)	Unpaid family worker (5)
Panel A: Main estimates					
Veiled × Post	0.012***	0.016***	0.002	-0.003***	-0.009***
	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)
Post	0.034***	-0.027***	0.048***	0.013***	0.009***
	(0.005)	(0.002)	(0.005)	(0.001)	(0.001)
Panel B: Trend-adjusted estimates					
Veiled × Post	0.013**	0.007**	0.003	-0.003	-0.001
	(0.006)	(0.003)	(0.006)	(0.002)	(0.003)
Post	0.022***	-0.025***	0.020***	0.006***	0.026***
(0.006)	(0.006)	(0.003)	(0.006)	(0.002)	(0.002)
Year fixed effects	No	No	No	No	No
Observations	933,810	933,810	933,810	933,810	933,810

Notes: Data are from the 2010-2017 HLFS. The sample includes all women aged 18-49. Panel A reports difference-in-differences estimates comparing labor market outcomes between predicted to be veiled and non-veiled women. Panel B includes $1\{\text{Veiled}\} \times t$, where t is a centered year variable ($t = \text{year} - 2010$). Each estimate include a dummy variable indicating whether the woman is native, dummies for the woman's education level, age, age squared, household size, and region fixed effects, with interaction terms between regions and *Post*. Standard errors (in parentheses) are clustered at bins of the predicted veiling probabilities for each woman. All estimates are weighted using cross-sectional survey weights from the wave in which the outcome was measured. ***, significant at the 1 percent level; **, significant at the 5 percent level; *, significant at the 10 percent level.

Table A5: *Effects of the 1997 education reform on women's religious practices*

Outcome	(1) Veiled	(2) Regularly pray	(3) Fast during Ramadan
Panel A: Static bandwidth			
Born after 1986	-0.058	0.046	-0.013
	(0.045)	(0.038)	(0.035)
Bandwidth	60	60	60
Covariates	No	No	No
Number of obs.	2,875	2,873	2,860
Mean	0.60	0.71	0.80
Panel B: Static bandwidth			
Born after 1986	-0.026	0.053	-0.002
	(0.044)	(0.036)	(0.034)
Bandwidth	60	60	60
Covariates	Yes	Yes	Yes
Number of obs.	2,786	2,784	2,772
Mean	0.58	0.70	0.83
Panel C: Donut-hole			
Born after 1986	-0.022	0.043	-0.001
	(0.047)	(0.037)	(0.034)
Bandwidth	60	60	60
Covariates	Yes	Yes	Yes
Number of obs.	2,633	2,631	2,620
Mean	0.58	0.70	0.83
Panel D: Optimal bandwidth CCT			
Born after 1986	-0.007	0.037	-0.024
	(0.048)	(0.031)	(0.030)
Bandwidth	49	82	74
Covariates	Yes	Yes	Yes
Number of obs.	2,277	3,840	3,440
Mean	0.58	0.70	0.83

Notes: The sample includes all women drawn from the 2013 DHS. Columns 1–3 report local linear RDD estimates for the probability of veiling, regular prayer, and fasting during Ramadan, respectively. All specifications include a linear function of month–year of birth and its interaction with the treatment indicator (born after January 1987). Panel A reports point estimates using a static bandwidth of 60 months, obtained with the Calonico-Cattaneo-Titiunik (CCT) algorithm proposed by Cattaneo et al. (2019), around the cutoff (January 1987). Panel B reports the same estimates while including background covariates: dummies for childhood place of residence, mother tongue, whether the woman's mother is literate, father's education level, the woman's month-of-birth fixed effects, and region-of-residence fixed effects. Panel C presents donut-hole RDD estimates, excluding women born within three months of the cutoff, and Panel D uses outcome-specific CCT optimal bandwidths. Standard errors are clustered at the month–year-of-birth level. ***, significant at the 1 percent level; **, significant at the 5 percent level; *, significant at the 10 percent level.

Table A6: *Effects of the 2002 civil code on women's religious practices*

Outcome	(1) Veiled	(2) Regularly pray	(3) Fast during Ramadan
Panel A: 1983–86 birth cohorts			
MonthExposure	-0.006	-0.003	-0.008***
	(0.004)	(0.004)	(0.003)
Number of obs.	1,144	1,143	1,138
Mean	0.60	0.71	0.84
Panel B: 1982–86 birth cohorts			
MonthExposure	-0.004	-0.004	-0.006**
	(0.003)	(0.003)	(0.003)
Number of obs.	1,452	1,451	1,445
Mean	0.60	0.72	0.84
Panel C: 1981–86 birth cohorts			
MonthExposure	-0.003	-0.003	-0.005**
	(0.003)	(0.003)	(0.002)
Number of obs.	1,754	1,753	1,746
Mean	0.60	0.72	0.84

Notes: The sample includes all women born between 1983 and 1986 in the 2013 DHS. Columns 1–3 report local linear RD estimates of the probability of veiling, regular prayer, and fasting during Ramadan, respectively, based on the dose–response design around the January 1985 cutoff. The running variable is the number of months from the cutoff, centered at January 1985, and interacted with an indicator for post-reform cohorts. Panel A reports point estimates using the 1983–1986 cohorts, Panel B expands the sample to 1982–1986 cohorts, and Panel C to 1981–1986 cohorts. All specifications include dummies for childhood place of residence, mother tongue, whether the woman's mother is literate, father's education level, the woman's month-of-birth fixed effects, and region-of-residence fixed effects. Standard errors are clustered at the month–year-of-birth level. ***, significant at the 1 percent level; **, at the 5 percent level; *, at the 10 percent level.

Table A7: *Robustness tests for labor market outcomes*

Outcome	Employment probability	Public sector employment	Private sector employment	Self-employed	Unpaid family worker
	(1)	(2)	(3)	(4)	(5)
Panel A: Excluding the 2013 wave					
Veiled × Post	0.010***	0.018***	0.001	-0.003***	-0.011***
	(0.004)	(0.002)	(0.003)	(0.001)	(0.002)
Veiled	-0.010***	-0.008***	-0.003	0.001	0.003
	(0.004)	(0.002)	(0.003)	(0.001)	(0.002)
Panel B: Controlling for removing the ban in universities					
Veiled × Post	0.012***	0.017***	0.002	-0.003***	-0.010***
	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)
After1991	0.005	0.034***	-0.028***	0.008***	-0.006**
	(0.005)	(0.002)	(0.004)	(0.001)	(0.003)
After1991 × Post	-0.001	-0.021***	0.011***	-0.001	0.006**
	(0.005)	(0.002)	(0.004)	(0.001)	(0.003)
Panel C: Controlling for the 1997 education reform					
Veiled × Post	0.010***	0.015***	0.001	-0.003***	-0.009***
	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)
After1986	0.008*	-0.034***	0.027***	-0.000	0.008***
	(0.005)	(0.002)	(0.004)	(0.001)	(0.002)
After1986 × Post	-0.018***	-0.003*	-0.018***	0.001	0.004**
	(0.003)	(0.002)	(0.003)	(0.001)	(0.002)
Panel D: Political alignment (Islamist party vote share)					
Veiled × Post × HighVoteShare	0.018***	0.007**	0.003	0.005**	-0.001
	(0.006)	(0.003)	(0.006)	(0.002)	(0.003)
Veiled × Post	0.004	0.013***	0.001	-0.006***	-0.009***
	(0.004)	(0.002)	(0.003)	(0.001)	(0.002)
Panel E: Using 2008 DHS for prediction and imputation					
Veiled × Post	0.010***	0.014***	-0.001	-0.002	-0.008***
	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)
Veiled	-0.016***	-0.007***	-0.005	0.001	-0.001
	(0.004)	(0.002)	(0.003)	(0.001)	(0.002)
Panel F: Alternative sample using traditional gender role attitudes					
Traditional × Post	-0.003	0.004***	-0.003	-0.002**	-0.004**
	(0.003)	(0.001)	(0.002)	(0.001)	(0.002)
Traditional	0.002	-0.001	0.003	0.000	0.002
	(0.003)	(0.001)	(0.002)	(0.001)	(0.002)
Wald test (p value)	0.000	0.000	0.191	0.620	0.002
Panel G: Alternative sample using information on praying					
Praying × Post	0.013***	0.003	0.011***	-0.001	-0.003*
	(0.003)	(0.002)	(0.003)	(0.001)	(0.002)
Praying	-0.010***	-0.003	-0.004	0.000	-0.002
	(0.003)	(0.002)	(0.003)	(0.001)	(0.002)
Wald test (p value)	0.734	0.000	0.010	0.084	0.000

Notes: Data are from the 2010–2017 HLFS. The sample includes all women aged 18–49. All regressions control for whether the woman is a native, education-level dummies, age, household size, and region and year fixed effects, with their interactions. Standard errors are clustered at the individual level. Data are weighted using cross-sectional weights from the wave in which the outcome was measured. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Table A8: *Effects on labor market outcomes using machine learning techniques*

Outcome	Employment probability (1)	Public sector employment (2)	Private sector employment (3)	Self-employed (4)	Unpaid family worker (5)
Panel A: Ensemble learning: SVM + RF + K-NN					
Veiled × Post	0.021***	0.024***	0.005**	-0.003***	-0.014***
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Veiled	-0.045***	-0.009***	-0.008***	0.001	-0.022***
	(0.003)	(0.001)	(0.002)	(0.001)	(0.001)
Panel B: Support Vector Machines					
Veiled × Post	0.024***	0.024***	0.006**	-0.002*	-0.013***
	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)
Veiled	-0.071***	-0.017***	-0.019***	0.001	-0.027***
	(0.005)	(0.002)	(0.004)	(0.001)	(0.002)
Panel C: Random Forest					
Veiled × Post	0.020***	0.024***	0.004	-0.003***	-0.014***
	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)
Veiled	-0.043***	-0.005**	-0.013***	-0.001	-0.018***
	(0.005)	(0.002)	(0.004)	(0.001)	(0.002)
Panel D: K-Nearest Neighbor					
Veiled × Post	0.015**	0.022***	0.003	-0.003**	-0.015***
	(0.006)	(0.005)	(0.003)	(0.001)	(0.002)
Veiled	-0.021	-0.013**	-0.003	0.001	-0.000
	(0.018)	(0.005)	(0.004)	(0.002)	(0.011)
Panel E: Neural Network					
PrVeiled × Post	0.040***	0.057***	0.003	-0.009***	-0.037***
	(0.005)	(0.003)	(0.004)	(0.002)	(0.002)
Veiled	-0.262***	-0.194***	-0.020***	-0.032***	-0.020***
	(0.009)	(0.004)	(0.008)	(0.004)	(0.005)
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	935,476	935,476	935,476	935,476	935,476

Notes: Data are from the 2010–2017 HLFS. The sample includes all women aged 18–49. Panel A reports difference-in-differences estimates using scarf predictions from ensemble learning methods; Panel B reports estimates using Support Vector Machines; Panel C reports estimates using Random Forests; Panel D reports estimates using K-Nearest Neighbor; and Panel E reports estimates using Neural Networks. The estimation follows Equation 1. Standard errors in Panels B–D are clustered by bins of predicted veiling probabilities, while Panel A reports robust standard errors. All estimates are weighted using cross-sectional survey weights from the wave in which the outcome was measured. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Table A9: *Effects on work hours, job quality, and occupations using machine learning techniques*

Outcome	Working hours (weekly)	Earnings (monthly)	Informal employment	Permanent job	Professionals	Technicians	Clerks
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Ensemble learning: SVM + RF + K-NN							
Veiled × Post	0.856***	6.985	-0.013***	0.032***	-0.000	0.015***	0.006***
	(0.116)	(5.125)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Veiled	-1.371***	4.949	-0.017***	-0.022***	0.004***	-0.009***	-0.005***
	(0.131)	(4.590)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Panel B: Support Vector Machines							
Veiled × Post	0.978***	6.213	-0.010***	0.033***	-0.001	0.015***	0.007***
	(0.157)	(6.860)	(0.002)	(0.003)	(0.002)	(0.001)	(0.001)
Veiled	-2.441***	-30.723***	-0.023***	-0.042***	-0.001	-0.009***	-0.007***
	(0.225)	(8.201)	(0.003)	(0.004)	(0.002)	(0.001)	(0.001)
Panel C: Random Forest							
Veiled × Post	0.763***	4.013	-0.010***	0.028***	-0.000	0.015***	0.006***
	(0.155)	(6.680)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
Veiled	-1.346***	1.906	-0.018***	-0.022***	0.005***	-0.007***	-0.003***
	(0.215)	(7.813)	(0.003)	(0.004)	(0.001)	(0.001)	(0.001)
Panel D: K-Nearest Neighbor							
Veiled × Post	0.611**	-0.768	-0.013***	0.025***	-0.001	0.014***	0.006***
	(0.263)	(5.941)	(0.002)	(0.005)	(0.002)	(0.003)	(0.001)
Veiled	-0.607	-2.740	-0.000	-0.019**	-0.001	-0.007***	-0.004**
	(0.628)	(29.779)	(0.009)	(0.009)	(0.002)	(0.002)	(0.001)
Panel E: Neural Network							
PrVeiled × Post	1.431***	-18.965	-0.030***	0.063***	-0.003	0.038***	0.015***
	(0.231)	(11.890)	(0.003)	(0.004)	(0.003)	(0.001)	(0.001)
Veiled	-8.747***	-709.706***	-0.040***	-0.215***	-0.097***	-0.036***	-0.044***
	(0.439)	(16.107)	(0.007)	(0.008)	(0.003)	(0.002)	(0.002)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	935,476	935,476	935,476	935,476	935,476	935,476	935,476

Notes: Data are from the 2010–2017 HLFS. The sample includes all women aged 18–49. Panel A reports difference-in-differences estimates using scarf predictions from ensemble learning methods; Panel B reports estimates using Support Vector Machines; Panel C reports estimates using Random Forests; and Panel D reports estimates using Neural Networks. The estimation follows Equation 1. Standard errors in Panels B–D are clustered by bins of predicted veiling probabilities, while Panel A reports robust standard errors. All estimates are weighted using cross-sectional survey weights from the wave in which the outcome was measured. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Table A10: *Heterogeneous effects by regions with a high share of veiled women ($\geq$ 50%)*

Outcome	Employment probability (1)	Public sector employment (2)	Private sector employment (3)	Self-employed (4)	Unpaid family worker (5)
Veiled × Post × MeanShare	0.007	0.005	0.002	-0.006**	0.003
	(0.007)	(0.004)	(0.006)	(0.003)	(0.003)
Veiled × Post	0.006	0.013***	0.001	0.002	-0.013***
	(0.006)	(0.003)	(0.005)	(0.002)	(0.003)
Observations	933,810	933,810	933,810	933,810	933,810

Notes: Data are drawn from the 2010–2017 HLFS and the sample includes all women aged 18–49. *MeanShare* is a binary indicator equal to one if the woman resides in a region where at least 50 percent of women are veiled (based on the 2013 DHS), and zero otherwise. The table reports difference-in-differences estimates comparing labor market outcomes between women predicted to be veiled and those predicted not to be veiled. All specifications include a dummy variable indicating whether the woman is native, dummies for the woman's education level, age, age squared, household size, and region and year fixed effects, along with their interactions. Standard errors (in parentheses) are clustered at bins of the predicted veiling probabilities for each woman. All estimates are weighted using cross-sectional survey weights from the wave in which the outcome was measured. ***, significant at the 1 percent level; **, significant at the 5 percent level; *, significant at the 10 percent level.

Table A11: *Worker efficiency in public sector before and after the headscarf removal*

Outcome	Match quality			Age	Earnings per hour
	Over qualified	Under qualified	Mismatch		
	(1)	(2)	(3)	(4)	(5)
Panel A: Main estimates					
Veiled × Post	-0.006	0.019***	0.013	-0.382	-0.022*
	(0.013)	(0.005)	(0.013)	(0.255)	(0.013)
Veiled	0.000	-0.011***	-0.011	0.651*	0.002
	(0.009)	(0.004)	(0.009)	(0.356)	(0.009)
Panel B: Trend-adjusted estimates					
Veiled × Post	-0.057**	0.003	-0.055**	-0.137	-0.055**
	(0.025)	(0.011)	(0.024)	(0.492)	(0.024)
Veiled	-0.021	-0.018***	-0.039***	0.752*	-0.012
	(0.014)	(0.005)	(0.014)	(0.427)	(0.013)
Observations	51,149	51,149	51,149	51,149	43,752
Mean Dep. Var.	0.281	0.024	0.306	34.23	2.835

Notes: Data are from the 2010-2017 HLFS. The sample includes all women aged 18-49. Panel A reports difference-in-differences estimates and Panel B includes $1\{\text{Veiled}\} \times t$, where t is a centered year variable ($t = \text{year} - 2010$). Each estimate includes a dummy variable indicating whether the woman is native, dummies for the woman's education level, age, age squared (except column 4), household size, and region and year fixed effects, along with their interactions. Standard errors (in parentheses) are clustered at bins of the predicted veiling probabilities for each woman. The last row reports the mean outcome for the sample. All estimates are weighted using cross-sectional survey weights from the wave in which the outcome was measured. ***, significant at the 1 percent level; **, significant at the 5 percent level; *, significant at the 10 percent level.

Table A12: *Effects of removing the headscarf ban on marriage probability*

Outcome	Ever married	Currently married	Divorced
	(1)	(2)	(3)
Panel A: Main estimates			
Veiled × Post	0.003	0.003	0.001
	(0.003)	(0.003)	(0.001)
Veiled	-0.002	0.001	-0.004***
	(0.005)	(0.005)	(0.001)
Panel B: Trend-adjusted estimates			
Veiled × Post	-0.030***	-0.034***	0.006**
	(0.007)	(0.008)	(0.003)
Veiled	-0.014**	-0.013*	-0.002
	(0.006)	(0.007)	(0.001)
Observations	933,810	933,810	933,810
Mean Dep. Var.	0.797	0.756	0.025

Notes: Data are from the 2010-2017 HLFS. The sample includes all women aged 18-49. Panel A reports difference-in-differences estimates comparing marriage and divorce probabilities between predicted to be veiled and non-veiled women. Panel B includes $1\{\text{Veiled}\} \times t$, where t is a centered year variable (t = year − 2010). Each estimate include a dummy variable indicating whether the woman is native, dummies for the woman's education level, age, age squared, household size, and region and year fixed effects, along with their interactions. Standard errors (in parentheses) are clustered at bins of the predicted veiling probabilities for each woman. The last row reports the mean outcome for the sample of veiled women before the policy change. All estimates are weighted using cross-sectional survey weights from the wave in which the outcome was measured. ***, significant at the 1 percent level; **, significant at the 5 percent level; *, significant at the 10 percent level.