Legal Barriers, Social Norms, and Female Labor Force Participation

Nishtha Sharma[*] Daniela M. Behr[†] Anna Fruttero[‡]

Abstract

Legal gender equality has advanced worldwide, yet female labor force participation remains uneven across countries. This paper studies why similar legal reforms generate different labor market outcomes across contexts. We develop a simple dynamic model in which women's labor market participation depends jointly on legal barriers and a social or normative cost that evolves endogenously with observed female labor force participation. The model formalizes a threshold mechanism: legal reform raises participation only if it is sufficiently strong to offset prevailing normative costs, which vary across societies. This framework clarifies why identical reforms may trigger self-reinforcing participation gains in some countries but fail to produce measurable effects in others. Guided by the model's testable predictions, we combine the Women, Business and the Law (WBL) index with gender-attitude measures from the World Values Survey and estimate the effects of large legal reforms using a staggered difference-in-differences design à la Callaway and Sant'Anna (2021) for 155 economies over 1990–2023. Consistent with the model, legal reforms increase female labor force participation only in countries with relatively progressive gender norms, while reforms have no statistically discernible effect in more norm-conservative settings. The results highlight that legal reform is necessary but not sufficient: its effectiveness depends critically on the normative environment in which it is implemented.

JEL Codes: J16, J21, L26, O17, K38, Z13

[*]CIOS, Faculty of Law, Charles University; nishthasharma.econ@gmail.com

[†]The World Bank; dbehr@worldbank.org

[‡]The World Bank; afruttero@worldbank.org

1 Introduction

Over the past five decades, governments worldwide have undertaken extensive legislative reforms to dismantle legal barriers to women's economic participation. As documented by the Women, Business and the Law (WBL) database, legal gender equality has improved steadily since 1970, with many countries expanding women's rights related to employment, entrepreneurship, property, and financial access (Hyland et al., 2020; World Bank, 2026). A large empirical literature shows that more gender-equal laws are associated with better economic outcomes for women, including higher labor force participation, improved access to finance, and faster aggregate growth (Christopherson Puh et al., 2022; Hyland et al., 2020; Tertilt et al., 2022).

Yet despite this progress, female labor force participation (FLFP) remains highly uneven across countries, even among those with broadly similar legal frameworks. In some contexts, legal reforms are followed by sustained increases in FLFP; in others, comparable reforms appear to have little observable effect. This discrepancy raises a fundamental question: why do similar legal reforms translate into markedly different labor market outcomes across countries?

The Europe and Central Asia (ECA) [1] region illustrates this puzzle clearly. Most ECA economies score relatively high on the WBL index, reflecting decades of reform aimed at expanding women's legal rights. Recent examples include the repeal of occupational restrictions for women in Kazakhstan in 2021 and similar reforms in Uzbekistan and Azerbaijan in 2022. Yet FLFP varies widely across the region. Countries such as Iceland and Moldova combine high legal equality with participation rates around 70 percent, while others, notably Türkiye, exhibit persistently low participation despite comparable legal scores. These patterns suggest that laws alone do not determine outcomes. The question has important macroeconomic implications: closing gender gaps in labor force participation could raise GDP by 15 to 20 percent in many economies, especially those where women are most constrained to work (Fiuratti et al., 2024; Goldberg et al., 2025; Pennings, 2022). A better understanding of the conditions under which such reforms succeed is essential to the design of effective policy.

A growing body of research points to social norms, defined as prevailing beliefs about

[1] In this paper, we follow World Bank geographical classification for the ECA region, comprising 46 economies: Albania, Armenia, Austria, Azerbaijan, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Hungary, Iceland, Ireland, Italy, Kazakhstan, Kyrgyz Republic, Latvia, Lithuania, Luxembourg, Moldova, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Russia, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Tajikistan, Turkey, Ukraine, United Kingdom, and Uzbekistan. Austria, Belgium, Greece, Italy, Luxembourg, and Norway, all geographically in ECA, are excluded from the analysis sample because their last major WBL reform ($\Delta WBL \geq 10$) occurred before 1990, leaving no observable pre-reform period in the FLFP data, which begin in 1990.

women's roles in the family and the workplace, as a separate and powerful determinant of women's labor supply. Gender norms shape expectations, social approval, and household behavior, and they tend to evolve slowly over time. At the same time, the literatures on laws and norms have developed largely in parallel, with limited attention to how legal reforms and social norms interact to determine participation outcomes (Gurbuz Cuneo et al., 2026; Hyland et al., 2020; Jayachandran, 2021).

This paper addresses this gap. We study how the effectiveness of legal reforms in raising female labor force participation depends on the normative environment in which they are implemented. Our central argument is that legal change operates through a behavioral channel that is inherently social: reforms lower formal barriers to women's work, but whether women take advantage of these new rights depends on the social cost of deviating from prevailing norms. Where such costs are low, reform can trigger a self-reinforcing process of entry; where they are high, reform alone may be insufficient to generate sustained change.

The empirical challenge is that legal change is endogenous, reforms are staggered across countries and time, and norms are difficult to measure comparably at scale. We develop a dynamic theoretical model of the interaction between laws and norms with a staggered difference-in-difference empirical design that estimates heterogeneous reform effects across normative contexts using norms measures from the World Values Survey.

We develop a dynamic model in which women decide whether to participate in economic activity by comparing individual benefits of participation against two costs: a legal barrier determined by the regulatory environment and a normative cost reflecting social disapproval of women's work. The normative cost evolves endogenously with observed female labor force participation, capturing the idea that participation becomes socially easier as more women work. The model delivers a sharp threshold implication: legal reform raises participation only if it is sufficiently strong to push the legal barrier below a level determined by prevailing norms and the distribution of individual returns to working. Identical reforms may therefore generate positive and persistent responses in progressive-norm contexts while producing little or no effect in norm-conservative settings. Hence, the intuition is straightforward: legal reform opens a door, but social norms determine whether women walk through it.

Guided by these testable predictions, we combine the WBL index with gender-attitude measures from the World Values Survey (WVS) and estimate the labor market effects of major legal reforms using a staggered difference-in-differences design à la Callaway and Sant'Anna (2021). Our analysis covers 155 economies over the period 1990–2023. To isolate meaningful legal shocks, we focus on large single-year improvements in the WBL index, which typically reflect comprehensive reforms affecting multiple legal domains.

The empirical results strongly support the model's implications. On average, legal reforms have modest and imprecise effects when pooling across all countries. Once we allow for heterogeneity by normative context, however, a clear pattern emerges. In countries with relatively progressive gender norms, legal reforms increase female labor force participation by about 1–2 percentage points within five years, with effects that grow over time. In countries with more biased norms, the same reforms have no statistically discernible impact. This divergence is particularly pronounced within the ECA region, where progressive-norm countries exhibit some of the largest post-reform gains in the sample, while biased-norm countries show no response.

These findings have important implications for both research and policy. They suggest that estimates of the average returns to legal reform may systematically understate the gains in settings where social norms are supportive, while overstating expected impacts where norms remain highly restrictive. More broadly, the results highlight that legal reform is necessary but not sufficient: its effectiveness depends critically on the normative environment in which it is embedded. The paper proceeds as follows. Section 2 reviews the related literature on laws, norms, and women's labor supply. Section 3 presents the conceptual framework and model. Section 4 describes the data and empirical strategy. Section 5 reports the main results and robustness checks. Section 6 discusses the implications for policy and future research.

2 Related Literature

This paper contributes to three strands of the literature: the effects of legal frameworks on women's economic outcomes, the role of social norms as determinants of female labor force participation, and the interaction between the two.

First, legal reform is widely regarded as a primary lever for expanding women's economic participation. Cross-country evidence documents positive associations between legal gender equality and female labor force participation (Hyland et al., 2020; Sever, 2023), finds that legal gender equality allows poorer countries to catch up with wealthier economies, promoting economic convergence (Sever, 2025), or shows that equal laws increase women's access to financial products (Demirgüç-Kunt et al., 2013; Perrin and Hyland, 2026). Further, structural models estimate that removing gender-based barriers could generate GDP gains of 15 to 20 percent in many economies (Chiplunkar and Goldberg, 2024; Gonzales et al., 2015; Pennings, 2022). Beyond the aggregate evidence on reforming gendered laws more broadly, a growing body of causal literature exploits variation in specific legal domains such as maternity leave policies, violence against women legislation, or labor laws to estimate effects on economic outcomes (for an overview: Saavedra Caballero et al. (2025)). For instance, Anukriti et al. (2025) apply a difference-in-difference design, finding that childcare laws increase female labor force participation by up to 2.2 percent. Using a difference-in-differences approach that exploits subnational spatial and cohort variation, Wilson (2022) studying the effect of child marriage finds positive effects on both schooling and the probability of employment. Assessing the spread of unilateral divorce laws in the United States Fernández and Wong (2014) find that married women's labor force participation increased significantly after changes in the divorce regime. This heterogeneity is further assessed in a systematic manner using the 35 areas covered in the WBL index. Fruttero et al. (2023) decompose the aggregate WBL index into its respective components and find that only nine out of the 35 areas are significantly associated with female labor force participation, with the largest effects concentrated among laws governing intra-household bargaining power, particularly equal rights to divorce and remarriage. Their analysis suggests that the relationship between legal reform and female labor force participation is not uniform across legal domains: some appear more strongly linked to labor market outcomes than others. The authors further note that governments should assess legal changes in light of country-specific norms. Our paper addresses a complementary source of heterogeneity: not which laws matter most, but under what normative conditions legal reform translates into participation gains. Where Fruttero et al. (2023) decompose variation across types of laws, we decompose variation across types of countries.

A second strand of literature establishes that norms or attitudes toward gender roles are powerful, persistent, and operate independently of legal and economic conditions. Following Bicchieri (2006) social norms can be understood as informal rules or behavior that persist because individuals prefer to conform when they expect others to do the same and believe that others expect conformity. In that sense, they differ from personal beliefs which may regulate behavior from an internal conviction, while social norms externally influence behavior of others (Bussolo et al., 2024).

This literature highlights the persistence of gender norms. Alesina et al. (2013) trace contemporary gender norms to historical plough agriculture; Fernández (2013) models how intergenerational learning about women's work generates the S-shaped trajectory of female participation in the United States; and Fernández and Fogli (2009) show that work and fertility decisions of second-generation American women are significantly predicted by cultural proxies from their country of ancestry, isolating beliefs from institutions. Jayachandran (2021) synthesizes this evidence, arguing that norms explain the large cross-country differences in female labor force participation that persist at comparable levels of development. More recently, Bussolo et al. (2024) show that persistent and in some cases increasingly conservative gender attitudes are closely linked to weak progress in female labor force participation in South Asia. A growing body of work also examines the channels through which norms constrain women's labor force participation. Norms reinforce traditional household divisions of labor (Bertrand et al., 2010; Folbre and Nelson, 2000) or shape perceptions of women's professional competence (Ridgeway and Correll, 2004). Complementing this perspective, Bursztyn et al. (2020) show that norms also constrain behavior in the present through perceived social expectations. In Saudi Arabia, men who privately support women working outside the home underestimate the extent of support among other men; correcting these misperceptions increases women's job search and labor-market engagement. Their findings highlight that norms operate not only through personal attitudes, but also through beliefs about what others consider socially acceptable. Our paper builds on this literature by showing that norms do not merely depress the level of FLFP but actively condition the dynamic response to legal change.

Third, despite growing evidence on the distinct effects of laws and norms on outcomes, relatively few studies systematically examine their interaction and the complementary reforms needed to make legal changes effective (for a detailed overview, see Gurbuz Cuneo et al. (2026)). Giuliano (2020) identifies this as an open question, noting that the feedback between formal institutions and cultural attitudes is theoretically ambiguous: legal reform may catalyze normative change, or entrenched beliefs may neutralize it. Existing work leaves this question largely unexplored in the context of legal reform. Fernández (2013) models endogenous cultural evolution but treats the institutional environment as exogenous. Hyland et al. (2020) and Fruttero et al. (2023) document the association between laws and outcomes but do not test whether normative context moderates reform effectiveness. A recent exception is Bussolo et al. (2025). The authors employ an instrumental variable approach using ancestral patriarchy as an instrument for contemporary norms, showing that conservative gender norms are a determinant of legal gender inequality. The authors show that more conservative societies adopt and implement less gender-equal laws, in part because such norms shape political selection and legislative behavior.

We bridge these strands of the literature by jointly studying leval reforms and social norms within a unified theoretica and empirical framework. On the theoretical side, we develop a simple dynamic model of FLFP in which legal barriers and prevailing normative costs jointly determine whether legal reform translates into participation gains. The model delivers a sharp threshold result: legal reform raises participation below a level that depends on the strength of social norms. In progressive-norm countries, the threshold is easily cleared and reforms translate into participation gains;

where norms remain strongly biased, reforms of the same nominal size may fail to trigger any response. We discuss extensions to incorporate explicitly endogenous norm dynamics and intrahousehold bargaining following Chiappori (1992), motivated by the finding in Fruttero et al. (2023) that laws governing women's household agency exhibit the strongest participation effects. These extensions would enrich the framework without altering its testable predictions. On the empirical side, we exploit staggered legal reforms across 155 countries using the Callaway and Sant'Anna (2021) estimator, estimating reform effects separately by normative context. The norm classification is based on a predetermined, cross-sectional country measure constructed from the World Values Survey, avoiding the identification concerns that would arise from interacting time-varying reforms with time-varying measures of norms. This design provides credible causal evidence that the effectiveness of legal gender reform is conditional on prevailing social norms.

We place special emphasis on the ECA region because it offers a bounded yet highly informative setting for studying this interaction. The region combines relatively high female educational attainment and a historically distinctive legacy of women's economic inclusion under state socialism (Brzozowska, 2015). At the same time, the region is far from culturally homogeneous, and gender attitudes vary substantially across countries (Cascella et al., 2024). This combination of relatively similar institutional backgrounds and human capital levels, together with pronounced variation in gender norms, makes ECA a particularly useful context for examining how identical legal reforms can generate divergent labor-market responses depending on the surrounding normative environment.

3 Conceptual Framework

This section develops a simple conceptual framework to clarify the mechanism through which legal reforms and social norms jointly shape women's labor force participation. The purpose of the model is not to provide a fully structural account of labor supply, but rather to formalize a key intuition that motivates the empirical analysis: the effectiveness of legal reform depends on whether it is sufficiently strong to overcome the prevailing social cost of women's economic participation.

We consider a dynamic setting in which women decide whether to engage in economic activity by comparing individual benefits of participation with two distinct costs: a legal barrier, reflecting formal restrictions imposed by laws and regulations, and a normative cost, reflecting social disapproval of women's work. A central feature of the framework is that the normative cost depends on observed FLFP. When few women participate, working deviates from prevailing expectations and carries a higher social penalty; as participation rises, this penalty declines. This feedback captures, in a reduced-form way, the idea that behavior both responds to and helps reshape social norms.

Setup

Let time be continuous, $t \in [0, \infty)$. A unit mass of women $i \in [0,1]$ each decide whether to participate in economic activity at time t. Let $w_t \in [0,1]$ denote female labor force participation (FLFP), the share of women who participate. Each woman faces two costs and one benefit:

- A *normative cost* $\gamma(1 - w_t)$, where $\gamma > 0$ captures sensitivity to prevailing gender norms. The factor $(1 - w_t)$ implies that participation is more costly when fewer women participate, since deviating from the observed pattern of behavior carries a larger social penalty.

- A *legal barrier cost* $c_l \in [0, \bar{c}]$, country-specific and constant across women within a period, capturing formal legal restrictions on women's economic participation.

- A *participation benefit* P_i, reflecting the perceived returns from working depending on women's individual characteristics such as education, ability, preferences and household dynamics. We assume $P_i \sim U[0, b]$ for tractability, where $b > 0$ is the upper bound of returns to participation.

A woman participates if her benefit exceeds her total cost:

$$P_i \ > \ \gamma(1 - w_t) + c_l. \tag{1}$$

Under $P_i \sim U[0, b]$, the probability that any given woman participates is

$$\Pr(\text{participation} \mid w_t) \ = \ \frac{b - \gamma(1 - w_t) - c_l}{b} \tag{2}$$

whenever the right-hand side is positive, and zero otherwise.

Dynamic Participation

Aggregate FLFP evolves as the difference between inflows (current non-participants who newly enter) and outflows (current participants who drop out):

$$\dot{w}_t \ = \ (1 - w_t) \cdot \Pr(\text{participation} \mid w_t) \ - \ w_t \cdot \left[1 - \Pr(\text{participation} \mid w_t)\right], \tag{3}$$

which, substituting Equation (2), gives

$$\dot{w}_t = (1-w_t)\cdot\frac{b-\gamma(1-w_t)-c_l}{b} - w_t\cdot\frac{\gamma(1-w_t)+c_l}{b}. \tag{4}$$

Setting $\dot{w}_t > 0$ and simplifying yields a clean threshold for participation to be expanding:

$$\dot{w}_t > 0 \quad\Longleftrightarrow\quad c_l < R(w_t) \equiv (1-w_t)(b-\gamma). \tag{5}$$

That is, FLFP grows at t if and only if the legal barrier c_l lies below the time-varying threshold $R(w_t)$. The threshold $R(w_t)$ is decreasing in w_t (the marginal incentive to enter weakens as more women have already entered) and decreasing in γ (a stronger normative cost lowers the legal barrier required to make further entry attractive).

Proposition 1 (Legal reform and norms)**.** *Consider an economy with current FLFP $w_t \in [0,1)$ subject to a legal reform that reduces the legal barrier from c_l^{pre} to $c_l^{post} = c_l^{pre} - \Delta c$, with $\Delta c > 0$.*

(a) ***Threshold for effective reform.*** *The reform raises FLFP, $\dot{w}_t > 0$, if and only if the post-reform legal barrier lies below the threshold:*

$$c_l^{post} < R(w_t) = (1-w_t)(b-\gamma).$$

Equivalently, the reform is effective if and only if Δc is large enough that $c_l^{pre} - \Delta c < R(w_t)$. Reforms that fail to clear the threshold leave $\dot{w}_t \leq 0$ and do not raise FLFP.

(b) ***Comparative statics on the threshold.*** *The threshold $R(w_t)$ is decreasing in γ, increasing in b, and decreasing in w_t:*

$$\frac{\partial R}{\partial \gamma} = -(1-w_t) < 0, \qquad \frac{\partial R}{\partial b} = (1-w_t) > 0, \qquad \frac{\partial R}{\partial w_t} = -(b-\gamma).$$

The same reform Δc is therefore more likely to clear the threshold (i) the weaker the normative cost (lower γ), (ii) the higher the upper bound on individual benefits (higher b, and (iii) the lower the current level of FLFP (provided $b > \gamma$).

(c) ***Knife-edge case.*** *If $\gamma \geq b$, then $R(w_t) = (1-w_t)(b-\gamma) \leq 0 \leq c_l^{post}$ for every $w_t \in [0,1]$ and every non-negative c_l^{post}. The threshold inequality is never satisfied: no reform of any size can make $\dot{w}_t > 0$. Legal reform alone cannot raise FLFP.*

A reform is effective if and only if it pushes c_l below the threshold $R(w_t)$. Three things shift the threshold: stronger normative costs γ lower it (making reform less likely to be effective), wider distributions (or average) of individual benefits b raise it (making reform more likely to be effective), and lower current participation w_t raises it (provided

$b > \gamma$). The knife-edge case $\gamma \geq b$ makes the threshold non-positive everywhere, so no positive c_l ever satisfies $c_l < R(w_t)$ and no reform can move FLFP at all.

We derive the following testable hypotheses.

Empirical Hypotheses.

H1. *Legal reforms increase women's participation.* Follows from Proposition 1(a): a large enough legal reform Δc that pushes c_l below $R(w_t)$ produces $\dot{w}_t > 0$.

H2. *Legal reforms persistently increase female labor force participation when norms are progressive (or norm conformity is low).* Follows from Proposition 1(b)–(c): higher γ lowers $R(w_t)$ at every w_t, so the same reform is less likely to clear the threshold; in the limit $\gamma \geq b$, no reform of any size is effective.

We do not explicitly endogenize the normative cost as a separate state variable. The factor $(1 - w_t)$ in our normative cost already captures the essential endogenous feature: as more women participate, the social penalty for participation falls, and this feedback is precisely what generates the threshold characterization in Proposition 1. A richer formulation would model n_t as a slow-moving state with its own law of motion. Such an extension would not, however, alter the *testable* predictions of the model. Our cross-country measure from the WVS is constructed as a country-level average across available waves rather than as a time series. Moreover, gender norms are widely documented to evolve slowly across generations (Alesina et al., 2013; Fernández, 2013; Giuliano, 2020).

We also abstract from intrahousehold bargaining for simplicity, as our identification strategy and data do not allow for separately identifying the bargaining channel. A natural extension along the lines of Chiappori (1992) would predict that legal reform raises women's outside options, amplifying the marginal effect of legal reform on participation. Our focus in this paper remains on how the normative environment shapes whether legal reform translates into higher participation.

4 Data and Empirical Design

Assessing our theoretical predictions, we combine several unique data sources to construct a country-year panel data set covering 155 economies over 1990–2023. The Women, Business and the Law database, FLFP drawing on estimates from the International Labour Organization (ILO), and individual-level data from the World Value Survey aggregated the country level.

Legal gender equality. We measure legal gender equality using the Women, Business and the Law (WBL) Index (World Bank, 2026). The index captures the extent to which laws and regulations affect women's economic opportunities over the course of their working lives, from entry into the labor force to retirement. The WBL Index aggregates 35 binary questions into eight dimensions: Mobility, Workplace, Pay, Marriage, Parenthood, Entrepreneurship, Assets, and Pension. It is scored on a scale from 0 to 100, where 100 indicates full legal parity with men. Topics included in the index are based on their associations with measures of women's economic empowerment and on evidence from economic literature, as discussed in Saavedra Caballero et al. (2025). The dataset relies on laws and regulations that are codified and currently in force. The WBL dataset provides annual observations for 190 economies over calendar years 1970–2023, enabling longitudinal comparison across the full 54-year span. While new measures on the implementation and enforcement of the laws were introduced in 2026, these data are currently not available for the same time-span (World Bank, 2026). Therefore, the focus of this paper is on women's de jure rights only.[2]

Figure 1 shows that over the past five decades, the global Women, Business and the Law score has significantly improved across all regions.

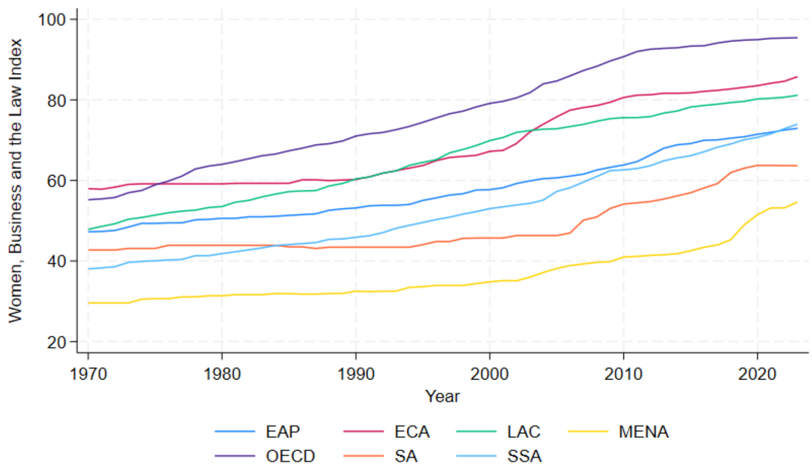

Figure 1: Women, Business and the Law Index, by Region

Female labor force participation. This variable captures female labor force participation, measured as the percentage of women ages 15 and older who are in the labor force. It is drawn from annual modeled estimates produced by the International Labour

[2]For a more detailed description of the database, the data collection and score calculation methodology as well as the scope of the eight indicators consult Hyland et al. (2020). Data set available at: https://wbl.worldbank.org/en/data/download-data

Organization (ILO) and available for nearly all countries from 1990 onward through the World Bank's World Development Indicators database, available for 174 countries (World Bank, 2025).

Gender norms. We measure prevailing gender attitudes using individual-level data from the World Values Survey (WVS) which gathers data on societal attitudes and values from respondents across numerous countries (Inglehart et al., 2022). The WVS includes a comprehensive range of questions concerning gender roles, perceptions, and expectations on a range of topics including beliefs about women's roles in the family, stereotypes about women's roles in leadership, or attitudes toward women's economic autonomy and role in the labor market (Gurbuz Cuneo et al., 2026). To construct the main variable for our analysis, we align our definition of biased social norms with the UNDP Gender Social Norms Index (GSNI) (United Nations Development Programme, 2023). Drawing on the WVS, we rely on the following seven statements to construct our gender norms index: (i) "A university education is more important for a boy than for a girl"; (ii) "Democracy means women have the same rights as men" ; (iii) "On the whole, men make better political leaders than women"; (iv) "When jobs are scarce, men should have more right to a job than women"; (v) "On the whole, men make better business executives than women"; (vi) "Is it justifiable for a man to beat his wife"; and (vii) "Is abortion justifiable". For each individual, we compute the share of these seven items on which the response reflects a gender-biased attitude following the same thresholds as the GSNI. We then average this individual-level bias share to the country level across all available WVS waves, yielding the *Biased Gender Norms Index*, interpreted as the proportion of gender-related attitudes that are biased in a given country. Our construction differs from the GSNI in aggregation: the GSNI classifies an individual as biased if they hold a biased view on at least one of the seven items, whereas we use the share of biased responses to preserve variation in the intensity of bias. We classify countries as having "Biased Norms" if their index exceeds the cross-country median, and we classify them as having "Progressive Norms" if their index falls below the cross-country median. The index is available for 89 countries in our analysis sample. Of the 40 ECA economies in the full panel, 36 have WVS data; the remaining 4 are not included in the norm-split specifications.[3]

[3]We also construct a measure of norm conformity from the WVS question on the importance of obedience as a quality to encourage in children (variable A196), averaged to the country level. This is available for 65 countries in our analysis sample and is moderately correlated with norm bias (r=0.54). We analyze the role of conformity as a moderator of reform effectiveness in Appendix Figure 10.

Controls. We control for Log GDP per capita from World Development Indicators in all our analyses and country fixed effects (by construction).

Dataset. The final data set contains 155 economies (5,267 country-year observations), including 40 countries in the Europe and Central Asia region. From the original 190 economies covered by the Women, Business and the Law dataset 16 economies are dropped because they lack female labor force participation coverage in the ILO modeled estimates over 1990–2023. An additional 19 economies are dropped because their only qualifying WBL reform occurred before 1990; these countries cannot contribute to the event-study estimator because the pre-treatment period is not observed. The resulting full panel of 155 economies is used to estimate the pooled effect of legal reform using the Callaway and Sant'Anna (2021) estimator. Our heterogeneity analysis by gender norms, which is the central contribution of the paper, is necessarily restricted to the subset of countries covered by the World Values Survey data on gender attitudes. This yields 89 economies (44 classified as progressive-norm and 45 as biased-norm based on median-split) for the norm-split specifications.

Table 1 provides descriptive statistics comparing ECA to the rest of the world: ECA has substantially higher WBL scores (79.8 vs. 60.0, $p<0.001$) and higher log GDP per capita (9.0 vs. 7.8, $p<0.001$), but only modestly and insignificantly higher female labor force participation (52.2 vs. 49.1, $p=0.15$). Average gender norm bias is similar across the two groups (0.36 in ECA vs. 0.37 in the rest of the world, $p=0.59$), although, as discussed above, ECA exhibits considerable within-region heterogeneity in normative attitudes. This gap between ECA's comparatively strong legal frameworks and its modest FLFP rates motivates our analysis.

We define the treatment *reform shock* as a single-year increase ≥ 10 points on the WBL index in order to map a continuous reform process into discrete reform episodes appropriate for staggered-adoption difference-in-differences. By design, the WBL index evolves both through major reforms and through smaller incremental updates. The applied 10-point threshold is intended to capture genuinely large and potentially behavior-changing legal reforms, rather than smaller legal updates that may be less likely to alter women's economic behavior. This cutoff isolates unique legal changes, approximately the top decile of positive annual index changes, that are more interpretable as salient shifts in the legal rights environment, as also required for reform effectiveness in theory (proposition 1a). Using smaller annual changes to define treatment would risk classifying incremental or more technical legal updates as *reform shocks*. Econometrically, the Callaway and Sant'Anna (2021) estimator is designed for discrete treatment adoption, making it important to identify a treatment onset that

Table 1: Descriptive Statistics: ECA vs Rest of World

Variable	Rest of World Mean	SD	N(RoW)	ECA Mean	SD	N(ECA)	p-value
WBL Index	60.04	(15.54)	115	79.80	(8.64)	40	0.000
Female LFP Rate (%, 15+)	49.10	(18.06)	115	52.21	(8.26)	40	0.147
Biased Gender Norms Index	0.37	(0.11)	53	0.36	(0.08)	36	0.589
Norm Sensitivity (Conformity)	4.51	(0.45)	39	4.34	(0.40)	26	0.111
Log GDP per capita	7.76	(1.39)	115	8.96	(1.30)	40	0.000
Ever Reformed (Δ WBL $\geq$ 10)	0.53	(0.50)	115	0.33	(0.47)	40	0.023
No. of Countries		115			40		

Notes: Country-level means and standard deviations on the 155-country analysis sample (1990–2023). The sample excludes 19 countries whose only qualifying WBL reform occurred before 1990 and who therefore cannot contribute to the event-study estimator. The Biased Gender Norms Index and Norm Conformity are constructed from the World Values Survey at the country level (averaged across all available waves); N is smaller for these rows because the WVS does not cover all countries. Combined coverage is 89 countries with bias data and 65 countries with conformity data. "Ever Reformed" indicates whether a country experienced any single-year increase in the WBL Index of 10 points or more during 1990–2023. p-values from Welch's t-test of difference in means.

corresponds to a meaningful break in the legal regime. Applying this threshold, 74 of the 155 economies included in the data set are treated, with 81 never-treated countries as controls. We show robustness to thresholds of 8 and 12 points in Table 2.

Examples for such legally disruptive *reform shocks* are legal reforms improving women's rights in multiple legal domains, typically reflected in adoptions of entirely new Civil or Labor Codes or comprehensive legal amdendments to multiple legal statutes. In 2004, for instance, the Government of Morocco made major amendments to its Family Code, known as the Moudawana, which covers personal status issues such as marriage, divorce, alimony, child support, child custody, and inheritance. At the same time, Morocco enacted a new Labor Code which modernized labor relations, specifically advancing women's rights by prohibiting workplace discrimination, criminalizing sexual harassment, and increasing paid maternity leave to 14 weeks. Combined, these legal changes resulted in an increase in the Women, Business and the Law score of 26.25 points. In 2021, Gabon, through Law No. 004/2021, implemented a major reform of Gabon's Civil Code aimed at equalizing spouses' rights within marriage and family life. The reform transformed Gabon's Civil Code from a husband-centered marital regime to a more gender-equal one, ending the husband's legal status as "head of family" and replacing it with joint management of the family by both spouses, providing spouses equal say over the family domicile, allowing each spouse the right to choose and exercise a profession, or opening a bank account in their own name without the other spouse's authorization. Together, these sweeping reforms resulted in an increase on the WBL index by 25 points. In ECA, Uzbekistan enacted significant reforms in 2023. It adopted a revised Labor Code that brought national legislation into closer alignment with international standards, including ILO Convention No. 100 on Equal Remuneration, by guaranteeing equal remuneration for work of equal value and lifting restrictions on women's employment in mining and other hazardous occupations. In the same year, Uzbekistan amended its Criminal Code and Code of Administrative Liability to explicitly criminalize domestic violence—including physical, psychological, and economic abuse within the family—and to prescribe sanctions for such conduct. These reform efforts increased the WBL score by almost 12 points.

Our main estimator is Callaway and Sant'Anna (2021), which constructs treatment effects separately for each cohort of countries (defined by the year of first reform) and each post-reform horizon, using only not-yet-treated and never-treated units as controls.[4] Each cohort-by-period treatment effect is estimated using the doubly-robust inverse probability weighting (DR-IPW) procedure of Sant'Anna and Zhao (2020), which combines an outcome regression for the conditional FLFP trajectory with an inverse-probability reweighting of controls based on a propensity score model for treatment timing (so that controls similar to treated units receive heavier weight). The doubly robust combination is consistent if either the outcome model or the propensity score model is correctly specified, providing a useful safeguard against misspecification of either component. We include log GDP per capita as a covariate and report event-study coefficients over a window of $k \in [-5, 5]$, with $k = -1$ as the reference period (the year immediately preceding reform).

To test the hypothesis, we estimate the model separately on countries with progressive norms (below-median WVS bias) and biased norms (above-median). The norm variable is cross-sectional and used only to sort countries into subsamples; it does not enter any regression and is therefore not absorbed by country fixed effects. As a secondary analysis, we further split by ECA vs. non-ECA to examine whether the norm-mediation mechanism explains ECA's FLFP rates.

5 Results

5.1 Average Reform Effect

Figure 2 presents the pooled Callaway and Sant'Anna (2021) estimates of the effect of reform shock on female labor force participation of the full sample of 155 economies. The

[4] We use the Callaway and Sant'Anna (2021) estimator rather than two-way fixed effects because reform timing is staggered across more than two decades with treatment effects that we expect to differ across normative environments, a setting in which TWFE event-study estimates are known to be biased by forbidden comparisons between later- and earlier-treated units (Goodman-Bacon, 2021; De Chaisemartin and d'Haultfoeuille, 2020; Borusyak et al., 2024).

overall ATT is 0.87 percentage points but is not statistically significant ($p = 0.19$), while the average post-reform effect is 0.47 percentage points and is marginally significant ($p = 0.06$). Examining the dynamic pattern, the effect is positive and statistically significant at the time of reform ($p = 0.03$), grows over subsequent years, and peaks at 0.77 percentage points four years after reform ($p = 0.04$), before becoming smaller and less precise by year 5.

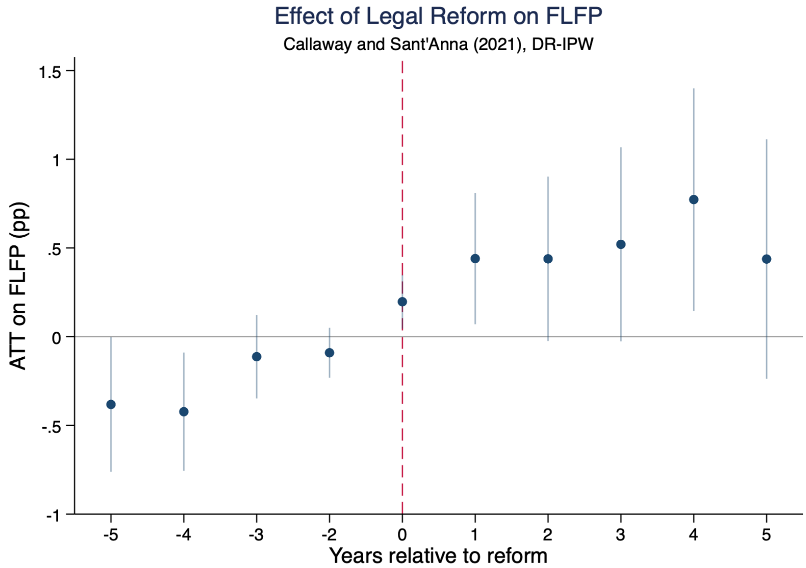

Figure 2: Effect of Legal Reform on FLFP (pooled sample)

Notes: Event-study estimates from Callaway and Sant'Anna (2021) using the doubly-robust IPW estimator with not-yet-treated units as controls. Coefficients show the ATT on female labor force participation (ages 15+, percentage points) at each year relative to a major legal reform (ΔWBL $\geq$ 10). The year before the reform or $t = -1$ is the omitted reference period, normalized to zero. 0 denotes the year of reform. Log GDP per capita included as a covariate. Vertical bars show 90% confidence intervals clustered at the country level.

The pre-period average is -0.32 pp ($p = 0.05$), at the edge of statistical significance and the weakest pre-trend in any specification we report. We therefore interpret the pooled estimates as suggestive evidence of a positive average reform effect that masks substantial heterogeneity across normative contexts. The norm-split specifications, which are the central contribution of the paper, satisfy parallel trends much more cleanly ($p = 0.15$ for progressive-norm countries and $p = 0.95$ for biased-norm countries; see next Section).

5.2 Heterogeneity by Gender Norms

The pooled estimates suggest that average effects may mask important cross-country heterogeneity. Figure 3 examines this directly by splitting the sample by our constructed

Biased Gender Norms Index dividing the sample into biased norm countries and progressive norm countries. The divergence between the two subsamples is striking.

In progressive-norm countries, reforms produce growing effects: +0.37 pp at $t = 0$ ($p = 0.010$), +0.48 pp at $t = 1$($p = 0.166$), +0.85 pp at $t = 2$ ($p = 0.035$), +1.31 pp at $t = 3$ ($p = 0.030$), +1.89 pp at $t = 4$ ($p = 0.009$), and +1.55 pp at $t = 5$ ($p = 0.049$). The post-reform average is +1.08 pp ($p = 0.011$). Pre-trends are cleanly satisfied on average ($p = 0.148$ on the pre-average). The growing trajectory is consistent with the virtuous cycle in Proposition 1(a) and Hypothesis H1: legal reform enables entry, which normalizes participation and weakens the normative cost, encouraging further entry.

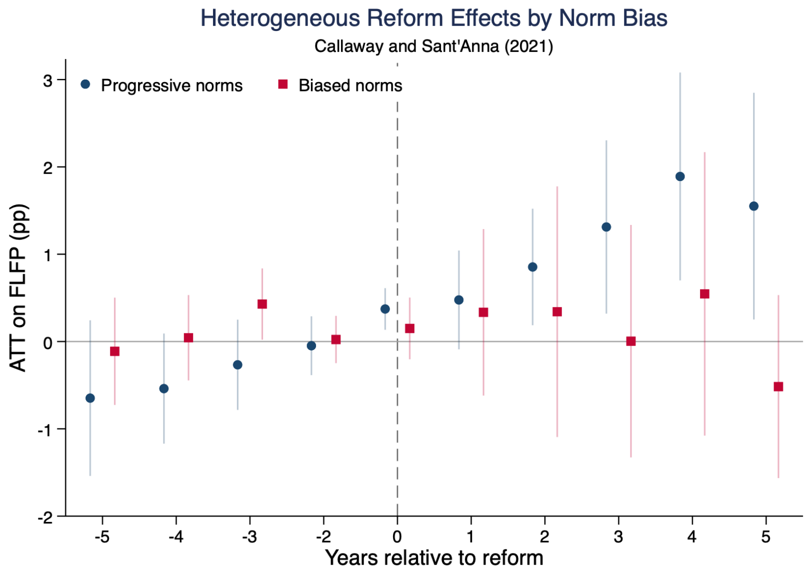

Figure 3: Reform Effects by Norm Bias

Notes: Event-study ATT estimates from Callaway and Sant'Anna (2021), estimated separately for countries with progressive gender norms (below-median norm bias) and biased norms (above-median). The year before the reform or $t = -1$ is the omitted reference period, normalized to zero. 0 denotes the year of reform (ΔWBL ≥ 10). Outcome is female labor force participation (ages 15+, percentage points). Log GDP per capita included as a covariate. Vertical bars show 90% confidence intervals clustered at the country level.

The divergence between the two subsamples grows over time. At $t = 4$, progressive-norm countries gain approximately 1.3 pp relative to biased-norm countries, widening to roughly 2.1 pp by $t = 5$. This evidence supports the hypothesis that legal reform and progressive norms are complements.

In addition, we find evidence that countries with low norm conformity experience a significantly positive effect of legal reforms on female labor force participation (post-reform average +0.95 pp, $p = 0.048$), while countries with high norm conformity show no effect (-0.05 pp, $p = 0.91$; Appendix Figure 10). The role of norm conformity further underscores how norms and culture determines whether legal gender equality translates into greater economic participation of women.

5.3 ECA vs. Rest of World

We begin by comparing reform effects in the Europe and Central Asia (ECA) region against the rest of the world. Figure 4 plots the event-study coefficients for the two groups. On average, the two regions look broadly similar: the Rest of World post-reform average is +0.46 pp ($p = 0.12$), and the ECA post-reform average is +0.53 pp ($p = 0.21$), with overlapping confidence intervals at every horizon. Neither region shows an individually significant post-reform coefficient at conventional levels, though effects build over time — reaching +0.74 pp in the Rest of World at $t = 4$ ($p = 0.10$) and +1.02 pp in ECA at the same horizon ($p = 0.10$). The Rest of World pre-average is marginally negative (−0.35, $p = 0.050$), while ECA pre-trends are clean (−0.27, $p = 0.58$). Taken at face value, this regional split offers no clear evidence that ECA is either especially responsive or especially unresponsive to legal reform.

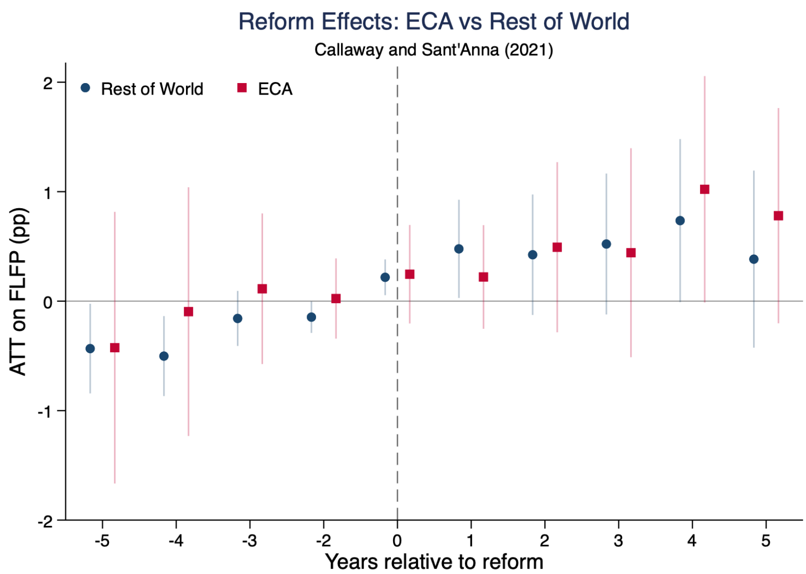

Figure 4: Reform Effects: ECA vs. Rest of World

Notes: Event-study ATT estimates from Callaway and Sant'Anna (2021) using the doubly-robust IPW estimator with not-yet-treated units as controls, estimated separately for countries in the Europe and Central Asia (ECA) region and the rest of the world. The year before the reform or $t = -1$ is the omitted reference period, normalized to zero. 0 denotes the year of reform (ΔWBL $\geq$ 10). Outcome is female labor force participation (ages 15+, percentage points). Log GDP per capita included as a covariate. Vertical bars show 90% confidence intervals clustered at the country level.

This pooled comparison, however, masks considerable heterogeneity within ECA. Figure 5 decomposes the ECA sample by norms. ECA with progressive norms shows the largest effects in any subsample: +0.64 pp at $t = 0$ ($p < 0.001$), growing to +1.52 pp at $t = 2$ ($p < 0.001$), +1.76 pp at $t = 3$ ($p = 0.001$), +2.69 pp at $t = 4$ ($p < 0.001$), and +2.14 pp at $t = 5$ ($p < 0.001$). The post-reform average is +1.46 pp ($p < 0.001$). Pre-trends are flat ($p = 0.76$ on the pre-average). ECA with biased norms shows no effect: the post-reform average is −0.12 pp ($p = 0.82$), no coefficient is individually significant, and pre-trends are clean ($p = 0.80$). This decomposition suggests that ECA's female labor force participation deficit does not reflect a failure of legal reform per se, but rather the composition of ECA countries across the normative spectrum. Reform gains are strong in the ECA countries where norms are progressive.

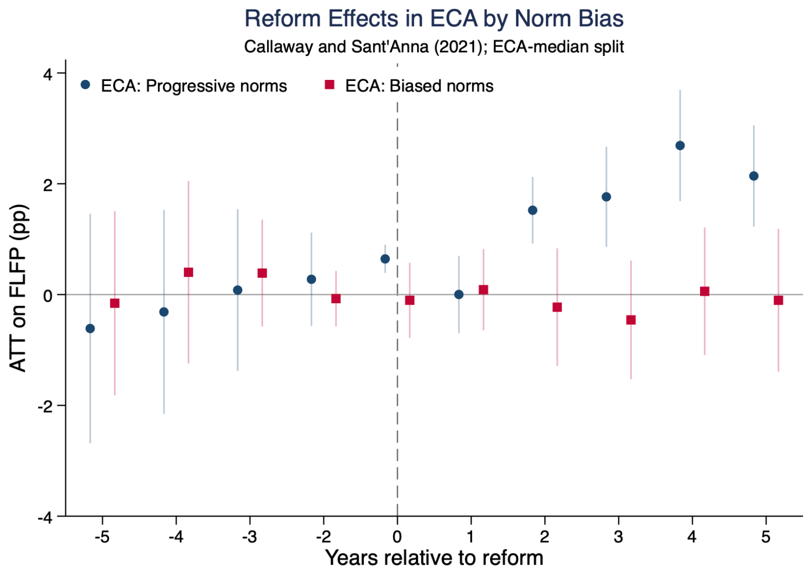

Figure 5: Reform Effects in ECA by Norm Bias

Notes: Event-study ATT estimates from Callaway and Sant'Anna (2021) using the doubly-robust IPW estimator with not-yet-treated units as controls, restricted to countries in the Europe and Central Asia (ECA) region and estimated separately by gender norm bias (progressive: below-median; biased: above-median). The year before the reform or $t = -1$ is the omitted reference period, normalized to zero. 0 denotes the year of reform (ΔWBL ≥ 10). Outcome is female labor force participation (ages 15+, percentage points). Log GDP per capita included as a covariate. Vertical bars show 90% confidence intervals clustered at the country level.

Outside ECA (Figure 6), the norm-based divergence is less clear. Neither progressive-norm nor biased-norm non-ECA countries show individually significant post-reform effects, and the confidence intervals are wide (progressive-norm Post-avg = +0.79, $p = 0.15$; biased-norm Post-avg = +0.26, $p = 0.79$). The progressive-norm non-ECA subsample exhibits a modestly negative pre-average (−0.87, $p = 0.07$), which warrants caution in causal interpretation. The sharper results within ECA may reflect the region's relatively homogeneous institutional environment, which provides a cleaner setting for isolating the moderating role of norms.

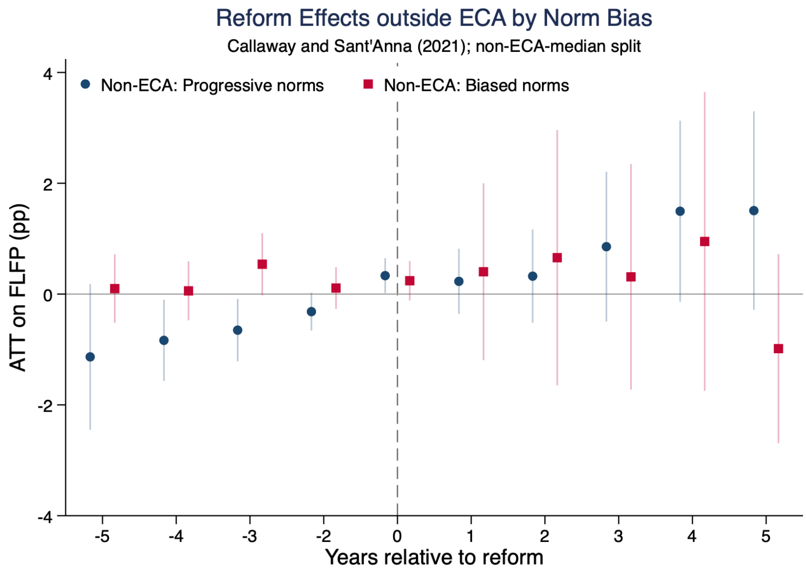

Figure 6: Reform Effects outside ECA by Norm Bias

Notes: Event-study ATT estimates from Callaway and Sant'Anna (2021) using the doubly-robust IPW estimator with not-yet-treated units as controls, restricted to countries outside the Europe and Central Asia (ECA) region and estimated separately by gender norm bias (progressive: below-median; biased: above-median). The year before the reform or $t = -1$ is the omitted reference period, normalized to zero. 0 denotes the year of reform (ΔWBL $\geq$ 10). Outcome is female labor force participation (ages 15+, percentage points). Log GDP per capita included as a covariate. Vertical bars show 90% confidence intervals clustered at the country level.

5.4 Robustness

We assess the robustness of the norm-based divergence through several additional specifications, all conducted separately on the progressive-norm and biased-norm subsamples to test whether the central pattern (significant effects in progressive-norm countries, null effects in biased-norm countries) holds beyond the baseline specification. The results are summarized in Table 2 and visualized in Appendix Figures 7–9.

Table 2: Summary of Callaway and Sant'Anna (2021) Estimates

Specification	ATT	Post Avg	p (Post)	p (Pre)	N_c
Full sample	0.870	0.468*	0.062	0.050	155
Main spec by norms (ΔWBL $\geq$ 10)					
Progressive norms	2.528**	1.076**	0.011	0.148	44
Biased norms	−1.459*	0.143	0.816	0.950	45
ECA by norms					
ECA, progressive	4.069**	1.403***	<0.001	0.809	16
ECA, biased	−0.092	−0.012	0.983	0.827	20
Conformity splits					
Low conformity	1.513	0.951**	0.048	0.609	33
High conformity	−0.671	−0.049	0.909	0.928	32
Single-reform countries by norms					
Progressive norms	2.617**	1.161***	0.008	0.219	42
Biased norms	−1.284	0.241	0.731	0.482	43
Threshold $\geq$ 8 by norms					
Progressive norms	2.189***	0.855**	0.016	0.051	44
Biased norms	−1.667*	−0.358	0.456	0.166	45
Threshold $\geq$ 12 by norms					
Progressive norms	1.595	0.530*	0.057	0.447	44
Biased norms	−0.769	0.771	0.305	0.394	45
Excluding Western Europe by norms					
Progressive norms	1.747	0.830*	0.069	0.348	36
Biased norms	−1.459*	0.143	0.816	0.950	45

Notes: Callaway and Sant'Anna (2021), DR-IPW, not-yet-treated controls, event window (−5,5). "ATT" is the overall simple average treatment effect on the treated; "Post Avg" is the average of post-reform event-study coefficients with p-value shown in the next column. "p (Pre)" is the p-value on the pre-treatment average. N_c = countries in the subsample. The Western Europe exclusion drops 12 additional economies from the progressive subsample (France, Germany, United Kingdom, Spain, Netherlands, Switzerland, Sweden, Denmark, Finland, Iceland, Ireland, Portugal); six other Western European countries (Austria, Belgium, Italy, Luxembourg, Norway, Greece) are already excluded from the main sample because their only qualifying reform occurred before 1990. The biased-norm row under "Excluding Western Europe" is identical to the main biased-norm specification because no biased-norm country in the sample is Western European. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Alternative reform thresholds. Appendix Figure 7 and Table 3 report the norm splits at thresholds of ΔWBL $\geq$ 8 and ΔWBL $\geq$ 12. The progressive-norm result is robust: the post-reform average is positive, statistically significant, and of similar magnitude across all three thresholds (+0.86 pp at $\geq$ 8, +1.08 pp at the main threshold of 10, and +0.53 pp at $\geq$ 12), with pre-trend averages not rejecting parallel paths on average. The biased-norm subsample also shows null or imprecise effects at every threshold. Thus, the qualitative divergence between the two subsamples does not depend on the specific 10-point cutoff used in the main specification.

Table 3: Norm Splits at Alternative Reform Thresholds

Threshold	Subsample	ATT	Post Avg	p (Post)	p (Pre Avg)
$\Delta WBL \geq 8$					
	Progressive norms	2.189***	0.855**	0.016	0.051
	Biased norms	−1.667*	−0.358	0.456	0.166
$\Delta WBL \geq 10$ *(main)*					
	Progressive norms	2.528**	1.076**	0.011	0.148
	Biased norms	−1.459*	0.143	0.816	0.950
$\Delta WBL \geq 12$					
	Progressive norms	1.595	0.530*	0.057	0.447
	Biased norms	−0.769	0.771	0.305	0.394

Notes: Callaway and Sant'Anna (2021) event-study estimates with doubly-robust IPW and not-yet-treated controls, window $(-5, 5)$. A reform is defined as a single-year increase in the WBL Index exceeding the stated threshold. ATT is the overall simple average treatment effect on the treated; Post Avg is the average of dynamic post-reform event-study coefficients. The final column reports the p-value for the test that the average of pre-treatment event-study coefficients equals zero. The progressive-norms post-reform average is positive and significant across all three thresholds. The biased-norms subsample yields null or imprecise estimates at every threshold. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Single-reform countries. Restricting the sample to the 58 countries that experienced exactly one major reform, a single-year change of 10 or more points toward gender equality, eliminates any confounding from subsequent legal changes. The norm-based divergence sharpens in this cleaner subsample: progressive-norm countries show a post-reform average of +1.16 pp ($p = 0.008$) and an overall ATT of +2.62 pp ($p = 0.016$), while biased-norm countries show a null effect (+0.24 pp, $p = 0.73$). The fact that the divergence is stronger when multi-reform countries are excluded suggests that the pooled estimates are, if anything, attenuated by noise from countries undergoing repeated legal changes (Appendix Figure 8).

Excluding Western Europe. A potential concern is that the progressive-norm result is driven by mature Western European economies with already-high participation. To address this, we re-estimate the norm splits after excluding Western European economies from the sample.[5] The progressive-norm post-reform average remains positive and borderline significant (+0.83 pp, $p = 0.069$), while the biased-norm subsample remains null (+0.14 pp, $p = 0.82$). The norm-based divergence is therefore not an artifact of

[5]The exclusion list comprises 18 countries (France, Germany, United Kingdom, Italy, Spain, Netherlands, Belgium, Austria, Switzerland, Sweden, Norway, Denmark, Finland, Iceland, Ireland, Luxembourg, Portugal, Greece), but six of these (Austria, Belgium, Greece, Italy, Luxembourg, Norway) are already excluded from the main analysis sample because their only qualifying WBL reform occurred before 1990. This robustness check therefore drops 12 additional economies.

Western European countries; it survives in a sample restricted to transition economies and the developing world, though precision is weaker (Appendix Figure 9).

6 Discussion

This paper shows that the effectiveness of legal reform depends critically on the normative environment in which it is implemented. xpanding women's formal economic rights does not generate uniform labor market responses across countries. Instead, the impact of legal reform is systematically conditioned by prevailing social norms. In countries where gender norms are relatively progressive, reform triggers a growing participation response over time; in countries where norms remain strongly biased, comparable reforms fail to produce sustained increases in female labor force participation.

The dynamic pattern observed in progressive-norm countries is consistent with the mechanism highlighted in the model. Legal reform facilitates initial entry into economic activity, thereby weakening the social penalty associated with women's work by shifting the descriptive norm. This feedback generates cumulative gains rather than a one-time adjustment: participation increases not as a discrete jump followed by stabilization, but through a gradual, self-reinforcing process. By contrast, in biased-norm countries, the initial behavioral response is too small to alter prevailing expectations, preventing the feedback mechanism from activating. The absence of dynamic amplification in these settings is consistent with the model's threshold characterization.

Importantly, the null results in biased-norm countries should not be interpreted as evidence that legal reform is irrelevant. Rather, they indicate that reform alone—absent changes in the surrounding normative environment—is insufficient to translate formal rights into widespread economic participation. This distinction matters for interpreting the existing literature: pooled estimates that average across normative contexts may understate the potential gains from reform where norms are supportive, while overstating expected impacts in contexts where social costs remain binding.

The estimated magnitudes are economically meaningful. In progressive-norm countries, the average post-reform effect is approximately 1.1 percentage points, with point estimates approaching 2 percentage points within four to five years of reform. These effects are substantially larger than those obtained in pooled specifications and underscore the importance of accounting for heterogeneity. Within the Europe and Central Asia region, the responses are larger still. Progressive-norm ECA countries experience average post-reform gains of about 1.4 percentage points, with participation increasing by more than 2 percentage points within five years. By contrast, ECA countries with biased norms exhibit virtually no response.

A natural concern is whether the heterogeneity we document reflects norms specifically or other country characteristics that correlate with the normative environment, such as state capacity, labor market structure, or education levels. Several features of our design mitigate this concern. The Callaway and Sant'Anna (2021) estimator constructs group-time average treatment effects using only not-yet-treated and never-treated units as controls. The doubly-robust inverse probability weighting specification, with log GDP per capita as a covariate, adjusts for residual differences in economic development between treated and control countries. The event-study framework further disciplines the analysis: pre-reform coefficients are flat and close to zero in both norm subsamples ($p = 0.15$ on the pre-average for progressive-norm countries and $p = 0.95$ for biased-norm countries), indicating that treated and control countries were on parallel participation trajectories prior to reform. The norm variable itself is cross-sectional and time-invariant, used only to partition the sample, not as a regressor, and therefore cannot be confounded by the within-country variation that identifies the treatment effects. The pattern survives a battery of robustness checks: restricting to the 58 single-reform countries, varying the reform threshold from $\Delta\text{WBL} \geq 8$ to ≥ 12, and excluding Western European economies from the sample. We cannot fully rule out that some unobserved confounder correlated with norms drives the results, but such a confounder would be unlikely to survive all of these specifications simultaneously.

The estimated magnitudes are economically meaningful. In progressive-norm countries, the average post-reform effect is approximately 1.1 percentage points, with point estimates approaching 2 percentage points within four to five years of reform. These effects are substantially larger than those obtained in pooled specifications and underscore the importance of accounting for heterogeneity. Within the Europe and Central Asia region, the responses are larger still. Progressive-norm ECA countries experience average post-reform gains of about 1.4 percentage points, with participation increasing by more than 2 percentage points within five years. By contrast, ECA countries with biased norms exhibit virtually no response.

The evidence on norm conformity further reinforces this interpretation. Countries characterized by lower conformity, where social expectations are less rigidly enforced, experience positive labor market responses to legal reform, while high-conformity countries do not. This suggests that it is not only the level of gender bias that matters, but also how strongly norms are socially enforced. From a behavioral perspective, this finding highlights the role of perceived social sanctions in shaping whether women exercise newly granted rights.

Several features of the empirical design increase confidence in this interpretation. The staggered adoption framework exploits within-country variation in reform timing and relies only on not-yet-treated and never-treated countries as controls. Pre-reform trends are flat in both normative subsamples, and the results are robust to alternative reform thresholds, single-reform samples, and the exclusion of Western European economies. While unobserved factors correlated with norms cannot be fully ruled out, any such confounder would need to explain the sharp divergence across subsamples and survive a wide range of robustness checks.

The policy implications follow directly from these findings. Legal reform remains a necessary foundation for expanding women's economic opportunities, but it is unlikely to be transformative on its own in settings where social norms impose high costs on women's labor force participation. In progressive-norm contexts, further legal improvements can be expected to generate cumulative gains. In norm-conservative settings, legal reforms may need to be complemented by interventions that directly target norms and their enforcement, such as investments in girls' education, public information campaigns that correct misperceptions of social support for women's work, increased visibility of female role models, or community-level engagement with men and traditional authorities.

More broadly, the results suggest that evaluations of legal gender reform should adopt longer time horizons and condition expectations explicitly on the normative environment. The central lesson is not that law is ineffective, but that law and norms interact. Durable gains in women's economic participation are most likely when formal legal change is accompanied—explicitly or implicitly—by social change that allows women to exercise the rights they formally hold.

References

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

Anukriti, S., Dinarte-Diaz, L., Montoya-Aguirre, M., and Sakhonchik, A. (2025). Filling the gaps: Childcare laws for women's economic empowerment. *AEA Papers and Proceedings*, 115:689–694.

Bertrand, M., Goldin, C., and Katz, L. F. (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, 2(3):228–255.

Bicchieri, C. (2006). *The Grammar of Society*. Cambridge University Press, New York, NY.

Borusyak, K., Jaravel, X., and Spiess, J. (2024). Revisiting event-study designs: robust and efficient estimation. *Review of Economic Studies*, 91(6):3253–3285.

Brzozowska, Z. (2015). Female education and fertility under state socialism in central and eastern europe. *Population (English Edition, 2002-)*, 70(4):689–725.

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

Bussolo, M., Ezebuihe, J. A., Muñoz Boudet, A. M., Poupakis, S., Rahman, T., and Sarma, N. (2024). Social norms and gender disparities with a focus on female labor force participation in south asia. *The World Bank Research Observer*, 39(1):124–158.

Bussolo, M., Rexer, J. M., and Hu, L. (2025). From patriarchy to policy: Norms, votes, and gender equal laws. World Bank, Policy Research Working Paper 11135.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Cascella, C., Pampaka, M., and Williams, J. (2024). Gender attitudes within and between european countries: Regional variations matter. *SAGE Open*.

Chiappori, P.-A. (1992). Collective labor supply and welfare. *Journal of Political Economy*, 100(3):437–467.

Chiplunkar, G. and Goldberg, P. K. (2024). Aggregate implications of barriers to female entrepreneurship. *Econometrica*, 92(6):1801–1835.

Christopherson Puh, K. M., Yiadom, A., Johnson, J., Fernando, F., Yazid, H., and Thiemann, C. (2022). Tackling legal impediments to women's economic empowerment. IMF Working Papers 2022/037, International Monetary Fund.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American economic review*, 110(9):2964–2996.

Demirgüç-Kunt, A., Klapper, L., and Singer, D. (2013). Financial inclusion and legal discrimination against women: Evidence from developing countries. World Bank, Policy Research Working Paper 6416.

Fernández, R. (2013). Cultural change as learning: The evolution of female labor force participation over a century. *American Economic Review*, 103(1):472–500.

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

Fernández, R. and Wong, J. (2014). Unilateral divorce, the decreasing gender gap, and married women's labor force participation. *American Economic Review*, 104(5):342–347.

Fiuratti, F. I., Pennings, S., and Torres Coronado, J. (2024). How large are the economic dividends from closing gender employment gaps in the Middle East and North Africa? World Bank, Policy Research Working Paper 10706.

Folbre, N. and Nelson, J. A. (2000). For love or money—or both? *Journal of Economic Perspectives*, 14(4):123–140.

Fruttero, A., Gomes, D. B., and Sharma, N. (2023). *Lawful progress: Unveiling the laws that reshape women's work decisions*. International Monetary Fund.

Giuliano, P. (2020). Gender and culture. *Oxford Review of Economic Policy*, 36(4):944–961.

Goldberg, P. K., Gottlieb, C. T. L., Lall, S., Mehta, M., Peters, M., and Ratan, A. L. (2025). The global gender distortions index (ggdi). Technical Report 34142, National Bureau of Economic Research.

Gonzales, M. C., Jain-Chandra, M. S., Kochhar, M. K., Newiak, M. M., and Zeinullayev, M. T. (2015). *Catalyst for change: Empowering women and tackling income inequality*. International Monetary Fund.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

Gurbuz Cuneo, A., Tribin Uribe, A. M., Trumbic, T., and Perrin, C. (2026). "I Want to Break Free": How laws and social norms open doors for women. World Bank, Policy Research Working Paper 11288.

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

Inglehart, R., Haerpfer, C., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E., and Puranen, B. (2022). World values survey: All rounds - country-pooled datafile. Dataset, Version 3.0.0.

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

Pennings, S. M. (2022). A gender employment gap index (gegi): A simple measure of the economic gains from closing gender employment gaps, with an application to the Pacific Islands. World Bank, Policy Research Working Paper 9942.

Perrin, C. and Hyland, M. (2026). Gendered laws and women's financial inclusion. *World Development*, 199:107234.

Ridgeway, C. L. and Correll, S. J. (2004). Unpacking the gender system: A theoretical perspective on gender beliefs and social relations. *Gender & Society*, 18(4):510–531.

Saavedra Caballero, F., Tribín, A. M., and Perrin, C. (2025). How laws promote economic opportunities: A review. World Bank, Policy Research Working Paper 11268.

Sant'Anna, P. H. and Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, 219(1):101–122.

Sever, C. (2023). Gendered laws and labour force participation. *Applied Economics Letters*, 30(19):2681–2687.

Sever, C. (2025). Legal gender equality as a catalyst for convergence. *Structural Change and Economic Dynamics*, 73:376–391.

Tertilt, M., Doepke, M., Hannusch, A., and Montenbruck, L. (2022). Marshall lecture 2022: The economics of women's rights. *Journal of the European Economic Association*, 20(6):2271–2316.

United Nations Development Programme (2023). 2023 Gender Social Norms Index (GSNI): Breaking down gender biases: Shifting social norms towards gender equality. Technical report, UNDP, Human Development Report Office, New York.

Wilson, N. (2022). Child marriage bans and female schooling and labor market outcomes: Evidence from natural experiments in 17 low- and middle-income countries. *American Economic Journal: Economic Policy*, 14(3):449–477.

World Bank (2025). World Development Indicators. Accessed December 5, 2025.

World Bank (2026). *Women, Business and the Law 2026: Benchmarking Laws for Jobs and Inclusive Growth*. World Bank, Washington, DC.

A Mathematical Proofs

Proof of Proposition 1

Starting from the law of motion in Equation (4), let $p(w_t) \equiv [b - \gamma(1 - w_t) - c_l]/b$ denote the participation probability at the current FLFP level. Then

$$\dot{w}_t = (1 - w_t)p(w_t) - w_t\big[1 - p(w_t)\big] = p(w_t) - w_t,$$

since the cross terms $(1 - w_t)p(w_t) + w_t p(w_t) = p(w_t)$ and the remaining $-w_t$ comes from the outflow expression. The sign of $\dot{w}_t$ is therefore the sign of $p(w_t) - w_t$:

$$\begin{aligned}
\dot{w}_t > 0 &\iff p(w_t) > w_t \\
&\iff \frac{b - \gamma(1 - w_t) - c_l}{b} > w_t \\
&\iff b - \gamma(1 - w_t) - c_l > b w_t \\
&\iff b - \gamma + \gamma w_t - c_l > b w_t \\
&\iff b - \gamma - c_l > w_t(b - \gamma) \\
&\iff c_l < (1 - w_t)(b - \gamma).
\end{aligned}$$

Defining $R(w_t) \equiv (1 - w_t)(b - \gamma)$, this proves the threshold characterization in part (a): $\dot{w}_t > 0$ if and only if $c_l < R(w_t)$. A reform from c_l^{pre} to $c_l^{\text{post}} = c_l^{\text{pre}} - \Delta c$ is therefore effective at w_t if and only if $c_l^{\text{post}} < R(w_t)$. Reforms that fail to clear the threshold leave $\dot{w}_t \le 0$ at w_t and do not raise FLFP. □

Part (b): comparative statics on $R(w_t)$. Differentiating $R(w_t) = (1 - w_t)(b - \gamma)$ directly,

$$\frac{\partial R}{\partial \gamma} = -(1 - w_t), \qquad \frac{\partial R}{\partial b} = (1 - w_t), \qquad \frac{\partial R}{\partial w_t} = -(b - \gamma).$$

For $w_t \in [0, 1)$, $(1 - w_t) > 0$, so $\partial R/\partial \gamma < 0$ and $\partial R/\partial b > 0$. The sign of $\partial R/\partial w_t$ is determined by $b - \gamma$; in the regime where reform can in principle be effective ($b > \gamma$, the negation of part (c)), $\partial R/\partial w_t < 0$.

These comparative statics imply that a reform of fixed size Δc is more likely to satisfy $c_l^{\text{post}} < R(w_t)$ when γ is smaller (the threshold is higher), when b is larger (the threshold is higher), and — provided $b > \gamma$ — when w_t is smaller (the threshold is higher). Under the uniform-distribution assumption $P_i \sim U[0, b]$, a shift in the support of the benefit distribution upward corresponds to an increase in b, and hence raises $R(w_t)$. □

Part (c): knife-edge case $\gamma \ge b$. If $\gamma \ge b$, then $b - \gamma \le 0$, so $R(w_t) = (1 - w_t)(b - \gamma) \le 0$ for every $w_t \in [0,1]$. Since any legal barrier satisfies $c_l \geq 0$, the inequality $c_l < R(w_t)$ is never satisfied, and $\dot{w}_t \leq 0$ for every w_t. Starting from any initial $w_0 \in [0,1]$, the trajectory is monotonically non-increasing and converges to $w^* = 0$. At $w^* = 0$, $p(0) = (b - \gamma - c_l)/b$, which under $\gamma \geq b$ is at most $-c_l/b \leq 0$; hence $\dot{w} = p(0) - 0 \leq 0$ at $w = 0$ and $w^* = 0$ is the unique steady state. A reduction in c_l does not change the regime $\gamma \geq b$, so $R(w_t) \leq 0$ continues to hold and the system remains pinned at zero. Legal reform has no effect. □

B Additional Figures

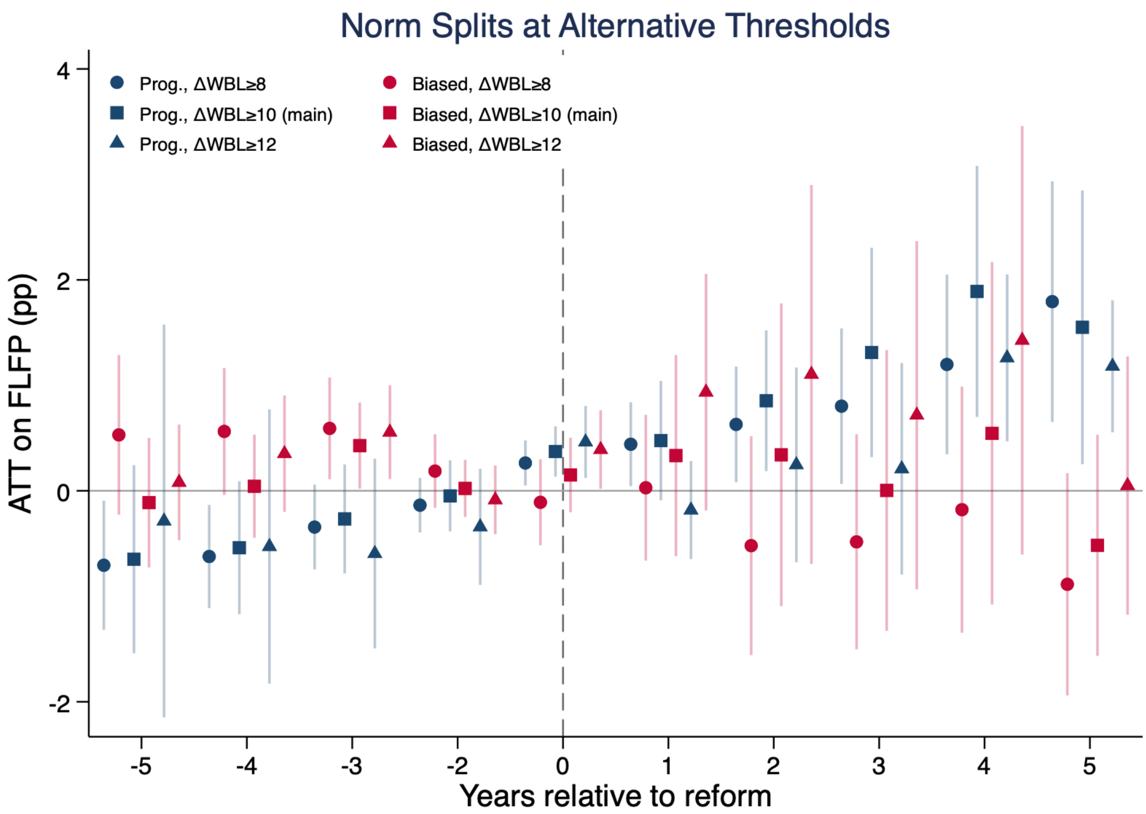

Figure 7: Alternative Thresholds

Notes: Robustness check varying the reform-intensity threshold used to define major legal reforms. Event-study ATT estimates from Callaway and Sant'Anna (2021) using the doubly-robust IPW estimator with not-yet-treated units as controls, estimated separately for countries with progressive gender norms (below-median norm bias) and biased norms (above-median) at three cutoffs: ΔWBL ≥ 8, ΔWBL ≥ 10 (main specification), and ΔWBL ≥ 12. The year before the reform or $t = -1$ is the omitted reference period, normalized to zero. 0 denotes the year of reform. Outcome is female labor force participation (ages 15+, percentage points). Log GDP per capita included as a covariate. Vertical bars show 90% confidence intervals clustered at the country level.

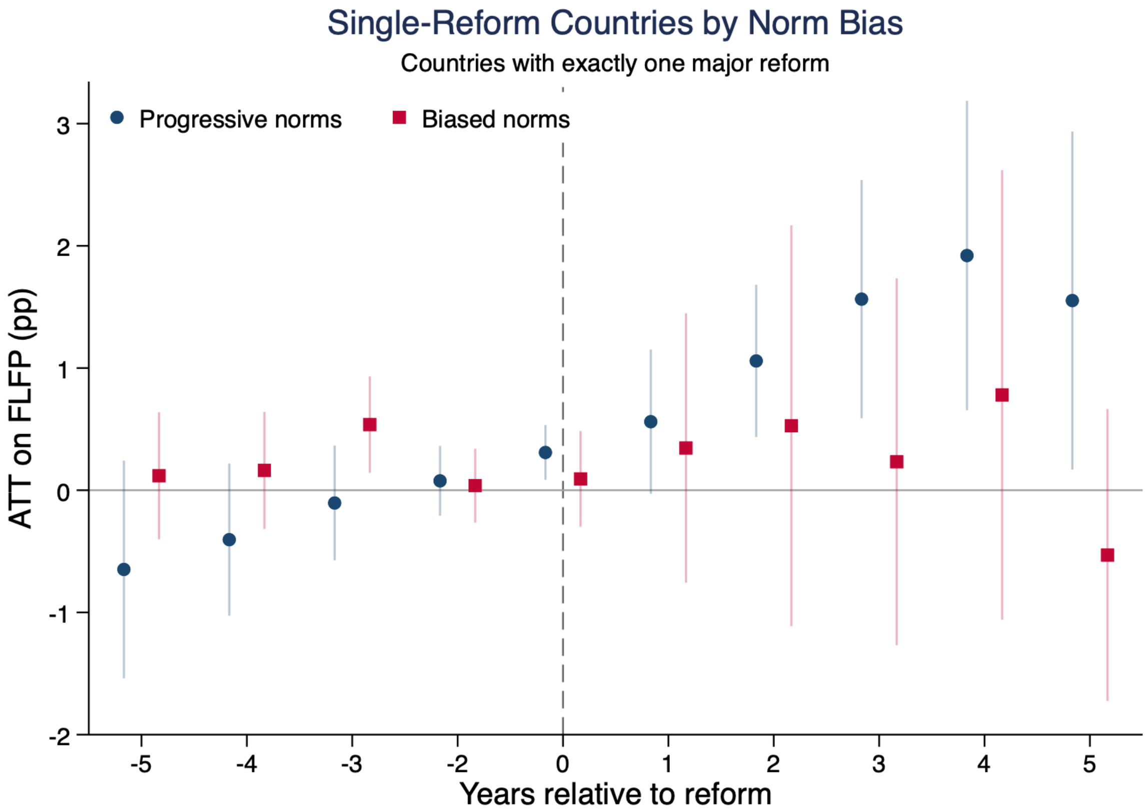

Figure 8: Single-Reform Countries Only

Notes: Robustness check restricting the treated sample to countries with exactly one major legal reform (ΔWBL $\geq$ 10) over the sample period, with never-treated countries retained as controls. Event-study ATT estimates from Callaway and Sant'Anna (2021) using the doubly-robust IPW estimator with not-yet-treated units as controls, estimated separately for countries with progressive gender norms (below-median norm bias) and biased norms (above-median). The year before the reform or $t = -1$ is the omitted reference period, normalized to zero. 0 denotes the year of reform. Outcome is female labor force participation (ages 15+, percentage points). Log GDP per capita included as a covariate with country fixed effects. Vertical bars show 90% confidence intervals clustered at the country level.

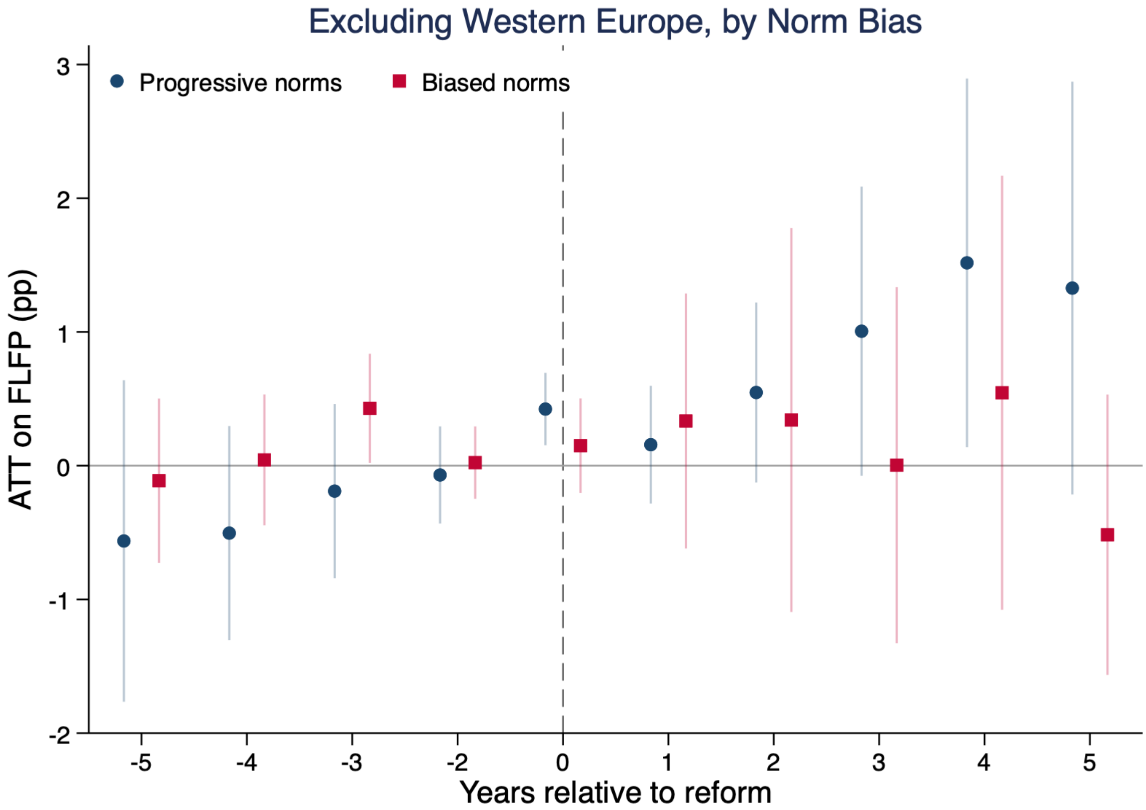

Figure 9: Reform Effects Excluding Western Europe, by Norm Bias

Notes: Robustness check excluding Western European countries (France, Germany, United Kingdom, Italy, Spain, Netherlands, Belgium, Austria, Switzerland, Sweden, Norway, Denmark, Finland, Iceland, Ireland, Luxembourg, Portugal, Greece) from the sample. Event-study ATT estimates from Callaway and Sant'Anna (2021) using the doubly-robust IPW estimator with not-yet-treated units as controls, estimated separately for countries with progressive gender norms (below-median norm bias) and biased norms (above-median). The year before the reform or $t = -1$ is the omitted reference period, normalized to zero. 0 denotes the year of reform (ΔWBL $\geq$ 10). Outcome is female labor force participation (ages 15+, percentage points). Log GDP per capita included as a covariate with country fixed effects. Vertical bars show 90% confidence intervals clustered at the country level.

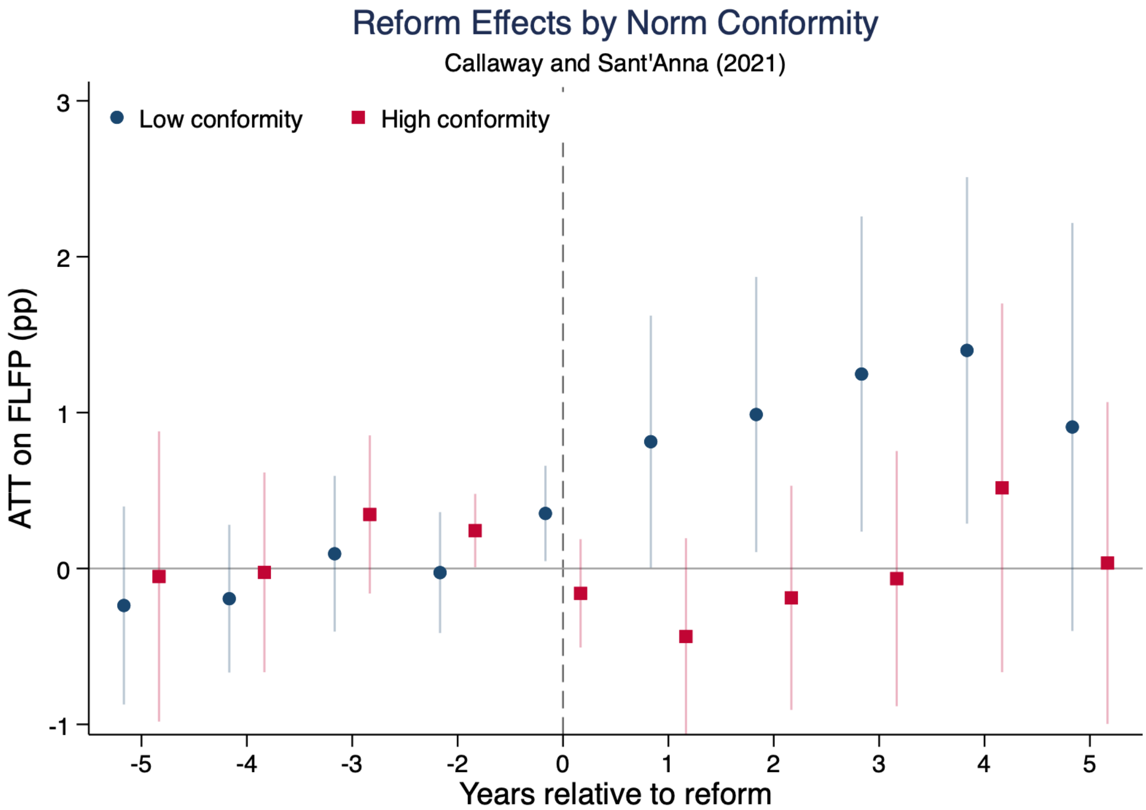

Figure 10: Reform Effects by Norm Conformity

Notes: Event-study ATT estimates from Callaway and Sant'Anna (2021) using the doubly-robust IPW estimator with not-yet-treated units as controls, estimated separately for countries with low and high norm conformity (split at the sample median). Norm conformity is measured as the country-year mean of World Values Survey item A196, which asks respondents to indicate whether obedience is an important quality to encourage in children. The year before the reform or $t = -1$ is the omitted reference period, normalized to zero. 0 denotes the year of reform (ΔWBL $\geq$ 10). Outcome is female labor force participation (ages 15+, percentage points). Log GDP per capita included as a covariate with country fixed effects. Vertical bars show 90% confidence intervals clustered at the country level.