Gender Diversity in Corporate Boards: Evidence from Quota-implied Discontinuities*

Olga Kuzmina†
New Economic School and CEPR

Valentina Melentyeva‡
ZEW and University of Mannheim

This version: April 2022

Abstract

We investigate the effects of women directors on firm value and operations, using data across European countries that introduced mandatory or voluntary regulation on female representation in corporate boards. We exploit quasi-random assignment induced by rounding, whenever percentage-based regulation applies to a small group. We find that having more women on board causally increases Tobin's Q and buy-and-hold returns. We further demonstrate that these positive effects are not explained by increased risk-taking or changes in board characteristics, but rather by reductions in empire-building activity. Our results highlight that promoting gender equality is aligned with shareholder interests.

Keywords: Gender diversity, gender quota, board of directors, firm performance

JEL codes: J16, J48, G34, G38, C18

* For their many suggestions, we would like to thank Ashwini Agrawal, Dmitry Arkhangelsky, Oriana Bandiera, Janis Berzins (discussant), Anna Bindler, Antonio Ciccone, Vicente Cunat, B. Espen Eckbo, Daniel Ferreira, Maria Guadalupe, Oğuzhan Karakaş, Egle Karmažiene (discussant), Adrien Matray, David Matsa, Marco Pagano, Daniel Paravisini, Steve Pischke, Sebastian Siegloch, Miriam Schwartz-Ziv (discussant), Vikrant Vig, as well as the seminar participants at the CEPR WE_ARE series, ECONtribute, the London School of Economics, University of Cambridge -- Judge Business School, EBRD, University of Bristol, University of Mannheim, City University of Hong Kong, Reichman University (IDC Herzliya), Bar-Ilan University, New Economic School, Higher School of Economics, and the participants of the 2020 American Economic Association meeting in San Diego, the 2020 European Economic Association Meeting, the 2020 German Economic Association Meeting, the 2020 Paris December Finance Meeting, the 2021 European Finance Association meeting, the 3rd CEPR Endless Summer conference.

† Corresponding author. Email: okuzmina@nes.ru; http://pages.nes.ru/okuzmina. Address: 3 Nobel St., Skolkovo, Moscow 121205, Russia.

‡ Email: vmelenty@mail.uni-mannheim.de; Address: L7 1, Mannheim 68161 Germany.

1 Introduction

Gender equality, and its main business-world facet – increasing female representation in boards of directors, – has become the agenda of policy makers across the world. "The Big Three" institutional investors (BlackRock, State Street, and Vanguard) also recently launched campaigns to promote gender diversity on corporate boards (Gormley, Gupta, Matsa, Mortal, and Yang, 2020). However, the seminal papers by Ahern and Dittmar (2012) and Matsa and Miller (2013) demonstrate that the effects of gender quota on firm value and short-term operating performance, respectively, are large and negative. These results raise a major policy dilemma: should gender equality be imposed at the expense of shareholders? In this paper, we use a novel discontinuity-based identification strategy to show that in the European countries that introduced percentage-based regulation, promoting gender equality is aligned with shareholder interests. This means that this policy dilemma does not exist, and institutional investors' money is smart. We further demonstrate that these positive effects mainly come from less empire-building activity on the part of women directors.

Estimating the causal impact of gender diversity on firm value is challenging. Early studies (Carter, Simkins, and Simpson, 2003) show a positive correlation between female directors and firm value, but Adams and Ferreira (2009) and Adams, Hermalin, and Weisbach (2010) identify several sources of endogeneity in the context of corporate boards. In a search for causal estimates, previous empirical work has largely relied on either the assumption that private firms are similar on unobservables to public firms (such as in Matsa and Miller, 2013), or that firms with different pre-quota shares of women are otherwise comparable to each other (such as in Ahern and Dittmar, 2012, and including most recent works by Hwang, Shivdasani, and Simintzi, 2019, and Eckbo, Nygaard, and Thorburn, 2021, among others).[1] But is the cure always better than the disease? Ferreira (2014) argues that some identification assumptions may be problematic.

We show empirically that firms that had more female directors at the time the regulation was announced had been growing faster prior to the regulation than firms with fewer women, in our sample of European public firms. We demonstrate that this implies that using past female share as

[1] While the first two papers use pre-quota share directly as the instrument, the latter two calculate the difference between the fraction of female directors required by the quota and that of the current board, and that also mechanially depends on how many female directors the firm already has.

part of the identifying variation would produce an overly negative estimate of the effect of female directors on firm value. To overcome this bias, we offer a new identification strategy, which is inspired by Angrist and Lavy's (1999) fuzzy-RD design, and exploits discontinuities that naturally arise due to rounding whenever a percentage requirement is applied to a small-sized group of people.

Specifically, we note that any percentage quota applied to a relatively small-sized group of individuals produces sharp increases in the actual minimum share of women that is to be achieved. This happens because women come in whole numbers. For example, with a quota of 25%, a board of size 4 has to have at least one woman, making it exactly a 25% as the minimum to be achieved. However, a board with 5 members has to have at least two women to satisfy the same 25%-requirement, achieving as much as 40%. Such a sizeable difference from what the quota prescribes means that firms with 5 board members will respond disproportionately more than firms with 4 members to the same percentage regulation, purely due to rounding, as long as board size is not perfectly flexible.

We consider board sizes measured prior to the exact minimum percentage is announced, hence it is not known to the firm ex ante which board sizes will eventually fall around this discontinuity. For example, for a 25% quota such close board sizes are 4 and 5, while for a 30% or 33% quota they are 6 and 7, and for a 40% quota – 5 and 6, etc. This ensures that the ex ante sorting of firms into the board sizes relevant for the specific percentage announced in a given country is likely to be close to random. Public firms within such narrowly-defined boards sizes form our close comparables. We further strengthen the argument by making the comparisons within a difference-in-difference framework, rather than a cross-section, so as to account for any time-invariant differences between boards of these slightly different sizes. Finally, we generalize this setting to multiple discontinuities within a country (i.e. also comparing 8- to 9-member and 12- to 13-member boards, in case of a 25% quota), to different percentages across countries (i.e. comparing 5- to 6-member boards, and 7- to 8-member boards, in countries where a 40% quota is introduced), and to different countries that introduced the same percentage in different years (i.e. comparing a 40% by 2015 mandate in Spain to a 40% by 2017 mandate in France). In our preferred and most saturated specifications, we can even identify the effects out of relative intensities, such as comparing the difference in performance between firms with 5 and 4 members within the same industry (which are predicted to have a 15% difference in minimum female share) to the difference in performance between firms with 9 and 8 members within the same industry (which are predicted to have only a 8.3% difference in minimum female share). The results are robust.

One limitation of our approach is that the ability to provide causal estimates under minimal possible identification assumptions comes at a cost of sample size requirements. Nevertheless, our final sample covers more than 60% of all BoardEx-Eikon public firms in the countries considered, and we are able to demonstrate that the main effects are similar for all countries together, as well as for many individual countries with different legal status and enforcement of quotas separately (such as the UK, France, the Netherlands, etc). Furthermore, the tight identification necessarily makes our causal effects local in nature. While we can convincingly demonstrate the marginal effects of essentially having one more woman in a board, they are intrinsically not designed to be extrapolated to much larger changes in board structures, where non-linearities would likely play a role.

We find a large and positive effect of more women in boards on Tobin's Q across European countries: a 5pp increase in the share of women increases Tobin's Q by 1 unit, which is about 0.6 within-firm standard deviations of this variable. This large effect, albeit non-extrapolable for much larger shifts in board composition, is consistent with boards achieving a "critical mass" of women that is necessary to becoming more active (Schwartz-Ziv, 2017), as well as larger effects for instrument-specific compliers within the heterogeneous effects framework (Imbens and Angrist, 1994).

This positive effect is also present when we look at individual countries in which our instrument provides enough variation to have a significant first stage. For the subsample of UK firms we additionally analyze the ultimate measure of investor performance – the buy-and-hold returns, which turn out to be about 1.6-3.8% higher in the annually compounded equivalent for firms that end up with a higher female share due to rounding. We show that these higher returns are not accompanied by higher loadings on common risk factors.

To explore the mechanisms behind these effects, we first look at board composition and behavior. We find that average board quality, as measured by average age, experience, network size, qualifications, and independence does not change. However, we do see a slight increase in average board meetings attendance, consistent with previous literature (Adams and Ferreira, 2009). While board meetings attendance contributes to the overall effect on value, its magnitude is somewhat too small to explain it fully. We proceed by decomposing the overall effect on Tobin's Q into the market-value and book-value components, and observe that the increase in the former primarily drives the effect. As our further analysis shows, this is not an artefact of a change in capital structure, or an increase in dividend payout, but rather the result of less empire-building activity. Specifically, we show that all size-related variables (assets, sales, employment) rise by less for firms with more female directors. Furthermore, we observe that such firms are less likely to incur merger-related expenses and to invest in purchases of fixed assets. The lower increases in sales, however, drive some short-run decreases in ROA. Nevertheless, they are not accompanied by any decrease in profit margins or labor productivity and wages, implying that firms still make the same profit out of each unit sold and worker quality does not deteriorate. Taken together and recalling that the long-run market reaction is positive, the evidence suggests that firms with more women engage less in inefficient operations and empire-building. In the context of overcoming the agency costs that lead to empire-building, these findings are also in the spirit of women being tougher monitors as in Adams and Ferreira (2009). This interpretation is also consistent with the literature that finds that women are less overconfident than men (see Croson and Gneezy, 2009, for a survey) and that overconfidence leads to investment distortions and empire building (Malmendier and Tate, 2005).

Our main contribution is twofold. Methodologically, we offer an empirical approach that enables studying the effects of any universal percentage-based regulation, under the minimal possible identification assumptions. Our identification strategy has its highest power when applied to relatively small-team settings, and as such it can be used in many settings outside of corporate finance and the context of gender. In political economy settings, for example, one could consider regulation applying to members of the Cabinet or members of the European parliament (but not e.g. members of Congress, which are too numerous to provide any meaningful discontinuities). It is also adaptable to other empirical setups, whenever close counterfactuals are of interest (such as in event studies).

Substantively, we apply this strategy to revisit the main empirical results from the literature on gender diversity in corporate boards, and show that the "common wisdom" of large negative value effects of female directors reverses. Our paper contributes to the literature studying the quota-imposed effects of gender diversity on firm performance. Besides seminal papers by Ahern and Dittmar (2012) and Matsa and Miller (2013), the recent evidence on Norway shows more mixed results: Eckbo et al. (2021) find zero effects, while Nygaard (2011) finds heterogeneous effects depending on information asymmetry. For Italy, Ferrari, Ferraro, Profeta, and Pronzato (2016) find no differences in performance, but some positive effects on stock prices. For a sample of 3 European countries, Comi, Grasseni, Origo, and Pagani (2020) find negative effects on productivity. The first US-based studies by Hwang et al. (2019), Greene, Intintoli, and Kahle (2020), and Meyerinck, Niessen-Ruenzi, Schmid, Solomon (2020) find negative market reaction to the introduction of gender quota in a sample of Californian firms, while Gertsberg et al. (2021) find evidence that this is explained by entrenched board dynamics rather than shareholders disliking the new female directors. Besides internal validity, our paper speaks towards external validity: rather than looking at one specific country or state, we consider virtually all (seven out of nine) European countries that introduced percentage quotas for public firms, both mandatory and voluntary. Finally, we propose a novel mechanism of female directors affecting firm value through less empire-building activity.

More broadly, our paper is part of a larger literature on corporate board structure, such as board diversity (see a survey by Ferreira, 2010), board representation (Jäger, Schoefer, and Heining, 2021), gender differences across directors (Adams and Funk, 2012), gender spillovers (Matsa and Miller, 2011), and gender and team performance in general (see some experimental evidence in Hoogendoorn, Oosterbeek, and van Praag, 2013). Our paper's final contribution to this broader literature is to highlight another implication of our empirical strategy. As we show, the mere existence of a significant first stage implies a specific average way of adjusting to the regulation: we show that firms mostly comply by switching male directors for female directors, rather than altering board size in a way to make it easier to comply with the regulation. Such a "sticky" board size implies that the costs of switching directors are perceived to be lower than those of altering board size, for the marginal firm. It is important because it implies that such percentage-based regulations may have additional unintended effects by shifting at least some firms away from their optimal board size.

The paper is organized as follows: Section 2 describes the empirical strategy; Section 3 discusses the data and provides summary statistics; Section 4 shows the first-stage results and validates the instrument; Section 5 presents the main results on the effects of female representation on value; Section 6 explores the mechanisms behind the effects; Section 7 concludes.

2 Empirical Strategy

To illustrate the idea behind our instrument, let's suppose that a firm faces a quota of 25% of women on the board (the specific number is used for illustration purposes, as it will apply to a majority of our sample). Does it mean that every firm that is *compliant* with this quota will have to have at least 25% women? Well, it turns out that most firms will actually have to have a percentage much higher than 25%, even if they want to only marginally comply with the 25%-quota. And the simple reason for that is that women (and men too) come in whole numbers. So a board of 2 directors will have to have at least 1 woman, making it a 50% share of women overall, while a board of 5 members will have to have at least 2 women, making it a 40% share of women. Only a board that is an exact multiple of 4 will have to have exactly 25% as the minimum to comply with the quota. Given how reluctant firms may have been in becoming compliant, even the differences in these minimum requirements induced by the same quota will likely produce enough powerful variation for us to identify the effects of interest. Overall, a firm with board size b, facing a quota q would need to have a minimum of

$$\frac{int((b-1)\cdot q)+1}{b},$$

where $int(a)$ is the integer part of a real number a, making this minimum a sawtooth-like function of the board size, such as the one in Figure 1 (drawn for the 25% case for concreteness).[2]

This pattern produces some natural discontinuities in the minimum required share of women, which is what we use in conjunction with our instrument. Since firms of very different board sizes

[2] This can be equally spelled as $\frac{roundup(bq)}{b}$, where $roundup()$ is the upward-rounding function.

are likely to be fundamentally different, we want to isolate the closest possible comparisons. To do that we investigate only the upward parts of this sawtooth-like pattern (this is analogous to Angrist and Lavy's, 1999, "discontinuity sample", highlighted in red in Figure 1). While any of the neighboring board sizes would be close enough comparables in terms of minimizing omitted variable bias concerns, the treatment is highest precisely at these jumps. Thus, using the close board sizes at these jumps, essentially leads to the highest signal-to-noise ratio. It additionally allows us not to rely on any additional functional form assumptions and extrapolation on how female presence and our variables of interest depend on the board size itself.

It is also essential that we never use the contemporaneous board size when constructing the instrument (since it is likely to be endogenous), but rather the original board size that existed before the quota and, importantly, its exact percentage were announced. This ensures that it is not known to the firm ex ante which board sizes will eventually fall around the discontinuity of interest, implying that the pre-existing sorting of firms into the board sizes relevant for the specific percentage announced in their country is likely to be close to random.

Our simplest possible instrument in this framework, $Right_i$, is then the dummy that takes a value of 1 for firms that were located just to the right of the kink in the discontinuity sample (i.e. 5, 9, 13, etc. in the case of 25% quota), in the year before the quota was announced, and a value of 0 for firms located just to the left of this kink (i.e. 4, 8, 12, etc. in the case of a 25% quota), and missing for all other values.[3] In the discontinuity sample, $Right_i$ also measures having one more woman (which is why the kink in share occurs to start with). Because both the minimum share and the absolute number of women jump at this kink, we have to be agnostic on the functional form of the relationship between women in boards and various dependent variables.[4]

[3] This most intuitive instrument has a much less intuitive mathematical formula that we only provide here for completeness:

$$Right_i = \begin{Bmatrix} 0 \text{ if } \frac{int((b_{i0}-1)\cdot q_c)+1}{b_{i0}} < \frac{int((b_{i0})\cdot q_c)+1}{b_{i0}+1} \\ 1 \text{ if } \frac{int((b_{i0}-2)\cdot q_c)+1}{b_{i0}-1} < \frac{int((b_{i0}-1)\cdot q_c)+1}{b_{i0}} \end{Bmatrix},$$

where b_{i0} is the board size of firm i in the year before the quota was announced (this year is country-specific), and q_c is the country-specific quota.

[4] Whether it is the share or the number of women that is more important is also unclear from the theoretical point of view. While some papers discuss the "critical mass" theory (e.g. Kramer et al, 2007), in the absence of separate exogenous shocks at each and every point of the schedule, we abstain from any functional form extrapolation. Nevertheless, for all of our kinks, being on the right also corresponds to having at least 3 women on the board (or at least 35% of women for smaller boards), which is often interpreted precisely as achieving a critical mass (see e.g.

To give an example of the identifying variation, let's consider for simplicity just one country, e.g. the United Kingdom (which is where most of our observations will come from anyway), which in 2011 published a recommendation by Lord Davies (2011) to incentivize larger firms to have at least 25% of women on boards by 2015. Our discontinuity sample in the UK will thus consist of firms that in 2010 (a year before the announcement) had 4, 5, 8, 9, 12, 13, etc. board members (highlighted in red in Figure 1). We will be making all of our comparisons within each of these red pairs, and to do that, we use the kink-specific fixed effects, λ_{kc}, that capture separate intercepts for firms that have 4 and 5 board members, vs firms that have 8 or 9 board members, vs firms that have 12 or 13 board members, etc. Due to these fixed effects, we compare firms only within each kink, but not across. Hence, none of our results can be explained by potential differences across firms with larger vs smaller boards (see e.g. Yermack, 1996). It is also important to note that firms naturally sort into these original board sizes before the actual percentage of the quota gets revealed, which even further reduces any concerns for selection of firms into specific board sizes (e.g. multiples of 4 vs one more member). Our main argument will thus be that this pre-existing sorting of firms within a kink, e.g. into whether to have 8 or 9 board members (and conditional on other things that we control for later), is likely to be orthogonal to the unobserved component in the value equation. This will be further weakened in the more saturated specifications.

One might argue that firms to the right of the kink mechanically have one more board member within each bin (as 5>4 and 9>8), and this might have its own effect on the dependent variables even in the absence of any quota (hence violating the exclusion restriction). We therefore weaken the identification requirements further and move to a difference-in-differences setup, finally estimating the first stage specification as follows:

$$Share_{it} = \gamma Post_{ct} Right_i + \lambda_{kct} + \lambda_{sct} + \lambda_i + \omega_{it}, \tag{1}$$

where $Share_{it}$ is the proportion of women on the board of firm i in year t, $Post_{ct}$ is the country-specific dummy variable that takes the value of 1 for the years after compliance, and 0 for the years before announcement, $Right_i$ is the instrument as defined above, λ_{kct} are the kink-specific fixed

Schwartz-Ziv, 2017).

effects (described above and kept country- and also year-specific, so as to absorb any country-year variation as well), λ_{sct} are the industry-year fixed effects (also specific to the country)[5], λ_i are firm fixed effects, and ω_{it} is the error term. Including kink-specific fixed effects, λ_{kct}, ensures that we never compare firms that belong to different kinks (i.e. the comparison group for 4s in the UK are only 5s in the UK an not e.g. 9s in the UK; and for 6s in Italy are only 7s in Italy, but not e.g. 7s in Belgium, and so forth). This is a difference-in-differences specification, since $Post_{ct}$ and $Right_i$ are automatically included and absorbed by (industry-)country-year and firm fixed effects, respectively.

This move to difference-in-differences helps to address potential pre-existing differences in the value of the company or other dependent variables for boards of different sizes. Additionally, it allows to absorb any non-linear relationship between $Share_{it}$ and board size that may exist even in the absence of a quota, under the assumption that the form of this non-linear relationship is similar before and after the reform. It is worth noting that our setup is different from the usual use of DID (with or without IV) in the quota setup (as in Ahern and Dittmar, 2012, and Matsa and Miller, 2013) in at least two important dimensions. First, the way how we construct counterfactual firms is different: we consider firms with very close ex ante board sizes, rather than firms with different ex ante shares of women or public and private companies, as these other papers. And second, because we have a meaningful and observable first stage, we don't have to make any assumptions on when the shock happens.[6] In our setup we can first empirically explore the dynamics of the first stage and measure when firms start responding to the instrument, and only then consider the second stage only where the instrument provides a powerful enough variation. This optimal selection of instruments (from the first-stage point of view) is possible as long as identification assumptions are maintained (see e.g. Paravisini et al, 2014, among others). It does not increase the measured impact of the quota on the dependent variables, because they are second-stage (Wald) estimates and change proportionally, irrespective of the size of the first-stage coefficient. As we will see further

[5]These are not necessary for identification and do not affect first-stage results. We add them in all first-stage specifications since they will be included as controls in the second stage for our dependent variables (Tobin's Q and others).

[6]This is perhaps one reason why different authors have disagreed on the timing of the Norwegian shock. While the share of women steadily rose from 2001 to 2009, papers have employed 2002 as the cutoff year (Eckbo et al, 2021), 2003 (Ahern and Dittmar, 2012), 2004 (Bertrand et al, 2019), 2005 (Nygaard, 2011), and 2006 (Matsa and Miller, 2013).

in Section 4.1, empirically firms on average do not respond sufficiently strongly to the instrument during the middle years (before mandated compliance year), so we cannot use them in the main part of the analysis. On the other hand, the differences in $Share_{it}$ between firms on the right and eft of the threshold start kicking in significantly after the mandated compliance year, provided by the regulation/recommendation in a given country. We therefore set $Post_{ct}$ to be 1 for all observations during the years post-compliance, and to 0 – during the years before announcement.[7]

The implicit assumption in the first-stage equation (1) above is that the effect on $Share$ of being to the right of the kink is the same at different kinks. This is fine, as long as we believe that that the effect mostly comes from having one more woman (rather than the percentage share itself), and it is constant across kinks. But if not, then we want to exploit the relative sizes of these jumps as well. We therefore proceed to defining our second instrument, which was already graphed in Figure 1 for the UK, in the following way:[8]

$$MinShare_i = \frac{int((b_{i0} - 1) \cdot q_c) + 1}{b_{i0}},$$

This allows us to proceed to our fullest specifications, where we estimate the first stage of our main equations of interest as follows:

$$Share_{it} = \gamma Post_{ct} MinShare_i + \lambda_{kct} + \lambda_{sct} + \lambda_i + \nu_{it}. \tag{2}$$

In essence, we want to additionally exploit the fact that the minimum share $MinShare_i$ is disproportionately larger between firms with 5 and 4 board members (40%-25%=15% difference) compared to that between firms with 9 and 8 board members (33.3%-25%=8.3% difference), and as such $Share_{it}$ is also expected to rise more on average in the former case than in the latter. This presents a very tight identification, under the minimal assumptions that are ever possible in the setting of a universal percentage quota.

As we will be measuring the effects over time, we cluster errors at the firm level to account for

[7] The actual date of firm-specific compliance is of course endogeneous, and hence not used.

[8] Our instrument is different from the *Shortfall* instrument, introduced by Eckbo et al. (2021) in that *Shortfall* still uses the ex ante share of women on the board as part of its construction. We discuss why instruments based on ex ante shares of women are likely to produce overly negative effects on value in Section 4.2.3.

arbitrary autocorrelation within firms and heteroskedasticity. However, it is important to emphasize again that the identifying variation is mostly cross-sectional, which means that we do not have to explicitly rely on timings associated with the quota (speed of compliance, when to define pre vs post, etc.), which some authors (e.g. Ferreira, 2014, and Eckbo et al., 2021) have argued might present a problem in terms of coincidences with various macro events and the associated differential impact of these events across firms with different shares of women. This reinforces the importance of bringing the comparison firms as close as possible and then let the data tell us when the change happens. This is precisely what we do and estimate (1) and (2) on the data 3 years before the quota announcement ($Post_{ct} = 0$) and 3 years after the quota compliance year ($Post_{ct} = 1$), skipping the intermediary years altogether.

A natural question is why being on the right of the kink *before* the quota can at all predict the share of females after the quota when firms have so many different ways of adjusting board composition to satisfy the quota? For example, those 9s that really don't want to have 33.3%, can just reduce the board size to 8 to attain the required minimum of 25%. If all firms to the right adjust like that, then γ will be close to zero. Furthermore, if firms instead adjust only by adding new female members until quota is satisfied, then those on the right may even end up with a *lower* share than those on the left (e.g. among firms with zero females that adjust by adding new members only, 4s will need 2 extra women for an average of $2/6 = 33.3\%$ and 5s will also need 2 extra women for an average of $2/7 = 28.6\%$, so that the difference-in-difference coefficient is -4.76%).[9] Only if firms exchange males for females at least to some extent would γ be positive (as it is in the extreme: 40% for 5-member boards and 25% for 4-board members, DID is +15%). However, how firms really adjust is ultimately an empirical question, and ex ante our identification strategy does not assume anything about their behavior.

This means three important things for us. First, from a purely econometric side, all other ways of adjusting, except exchanges, will bring γ closer to zero (or even negative), reducing the power of the first stage, and making it harder for us to track any changes in $Share_{it}$ at all (and later in the

[9] This and further discussions on the expected magnitudes of the DID estimates implicitly assume that the average original share of women for firms on the right and left are the same. While this is never assumed in the identification itself, we do check that this still holds empirically (see footnote 12 and Appendix Figures 1 and 2).

dependent variables of interest). Second, if we do find a positive and significant γ (which we do), this means that the predominant way of adjusting to the quota is actually exchanging males for females. This is an important empirical observation about board size being so sticky that all other ways of adjusting are more costly, at least on the margin. And third, different countries may have different ways of adjusting due to a variety of institutional and cultural reasons, suggesting that if we were to consider individual countries one-by-one, we may find a different γ across countries. To sum up, our identification strategy does not assume that all firms adjust by exchanging males for females and/or that the board size is constant. Instead, we empirically explore whether this is on average true or not, and then use this empirical fact to track changes in the variables of interest.[10]

Finally, in principle we could also attempt to directly exploit the different tightness of the regulation within countries.[11] For example, in the UK the original 25%-recommendation of Lord Davies applied to FTSE100 firms, while FTSE 350 companies were supposed only to "set out the percentage of women they aim to have on their boards by 2015". Similarly, in France the quota applied to firms that meet certain turnover or employment criteria. Using such a narrow sample potentially has the advantage of increasing the power of the first stage if there is no measurement error in applying these criteria and if other firms, that are not subject to recommendation/quota, do not change their behavior. However, this approach of narrowing down the list of firms also has a disadvantage. When the rest of firms do in fact change their behavior (e.g. in the expectation that soon a similar regulation might apply to them too, or because they want to keep up with the larger firms), the power of the first stage becomes weaker if they are ignored. Empirically, we find that using all firms that may in principle react to the regulation, provides a stronger first stage.

Importantly, the fact that the nature of regulation in many countries is recommendatory, rather than mandatory, is irrelevant, because the exogenous "shock" is whether the firm had been on the right or left of the threshold before the regulation was announced. As such, the instrumental-

[10] There is a group of firms for which the first-stage γ is expected to be zero (and is also zero empirically) – firms which already had a higher share of women than is needed to satisfy the quota. Such firms present only 4% of the original discontinuity sample and we drop them throughout the paper to increase the power of the overall first stage. The results are, however, the same if these few firms are kept, and are available upon request.

[11] As for exploiting differences across countries, there are only 7 countries in our sample and they differ in far too many dimensions to make any sample split credibly attributable to one single difference across the two sets of countries. We discuss this further when we consider country-level results in Section 5.1.2.

variables strategy isolates only the exogenous part of the variation in the share of women that is driven by this shock only, rather than by any unobservables that may be relevant to voluntary compliance decisions. Our results are therefore informative about the effects of any type of regulation, recommendatory or mandatory, that creates such a shock in the recommended/mandated share of women solely due to a different board size. This is the most standard use of IV in the presence of non-perfect compliance (Imbens and Angrist, 1994, Angrist, Imbens, Rubin, 1996) that will estimate the LATE causal effect of exogenous changes of the share of women on firm performance and outcomes. We futher discuss the LATE interpretation in Section 5.1.1.

3 Data and Summary Statistics

The results of our paper are based on two sets of data. We use an unbalanced panel of listed companies across European countries from BoardEx to obtain the director-level information on gender, age, number of qualifications, network size, role (independent or not), and other characteristics, and average them at the firm level. We then merge this dataset with financial data on public companies from Eikon. The exact set of countries is comprised of the United Kingdom, France, Italy, Belgium, Spain, the Netherlands and Norway. This is the exhaustive list of European countries that introduced formal (through quotas) or informal (through advisory recommendations) regulations on gender diversity that satisfy the following criteria: 1) this regulation contains a specific minimum percentage that has to be achieved (otherwise, without clear targets we would not be able to exploit our discontinuity-based instrument); 2) it applies to non-state-owned publicly-listed firms; 3) it was announced earlier than 2017 (so that we have enough years of observations to measure the outcomes); and 4) there are at least 20 firms in the discontinuity sample (so that our multiple-fixed-effects specifications can be estimated).[12] The period of study varies depending on the country and the respective year when the regulation was introduced and covers all years from

[12] This is not a hard constraint per se, as it rules out only Iceland with its 3 firms in the discontinuity sample). We do, however, have to exclude Germany, because listed companies above 2,000 employees (precisely the ones subject to the 30% gender quota after 2016) have to have either 12, or 16, or 20 supervisory board seats, depending on the number of employees (Co-determination Act, 1976). As such, there are no comparable firms within any discontinuity bin.

3 years before quota/regulation first announcement to 3 years after and including the compliance year (or to 2019, whichever is earlier). The complete coverage of countries with a short description of regulation and the relevant years is presented in Table 1.

As expected due to BoardEx coverage, most of our sample (slightly less than 60%), comes from the United Kingdom. We will therefore present all the analysis both for the UK alone, as well as for all countries together. The counts in Table 1 show the number of firms in the discontinuity sample as of the year before the regulation announcement. For example, there were 445 public firms in the UK in 2010 that had board sizes that are either exact multiples of 4, or had one more board member. The second largest country is France with 135 firms in 2009 in the relevant discontinuity sample (which for a quota of 40% covers many more board sizes). On the other hand, there were only 20 firms in the relevant discontinuity sample in Norway and 29 in the Netherlands. Since we also perform our analysis for the UK alone, and our results are very similar, they are not driven by any of these countries having very small sets of firms.

In Figure 2 we further show the exact distribution of pre-announcement-year board sizes, by country, with red (dark) bars representing board sizes of the discontinuity sample, and the grey (light) bars representing all other boards sizes not used in the analysis, across BoardEx-Eikon listed firms. We also do not consider (almost mechanically) very small boards of fewer than 4 members in the year before the quota announcement.[13] Since French quota applies to non-executive members, the relevant discontinuity samples are based on the ex ante number of non-executive directors, rather than total board size. Depending on the exact quota percentage, which defines the board sizes to be included in the discontinuity sample, and the distribution of firms across board sizes, our sample covers from 54% of these firms in the UK, to 71% in France, and above 70% in most other countries, for a total of about 60% of all BoardEx-Eikon public firms in these countries. In the unreported results, we also show that variable distributions in the pre-announcement year are similar in the discontinuity sample and out of it, within each country, which is expected given the

[13] For some countries for smaller boards the quota is defined as of the minimum number of women to be achieved, and we account for that in the analysis. For example, in Norway boards of less than 3 people should have at least 1 woman, boards of 4-5 people – at least 2 women, boards of 6-8 people – at least 3 women, and larger boards – at least 40% of women. This schedule naturally defines the discontinuity samples to be the board sizes of 3 vs 4, 5 vs 6, and 8 vs 9. This small-board issue, however, is not relevant for every country: for example in Spain, which requires 40% of each gender, there are simply no firms with fewer than 9 board members in sample.

way it is constructed. This also speaks to the generalizability of our results to boards of different sizes. In what follows, we refer to the discontinuity sample as the sample.

Table 2 presents summary statistics for the main variables of interest, with all continuous variables winsorized at 1% tails. We present the statistics for all years used in the analysis. Companies in our sample have on average 21 bn Euro total assets (0.5 bn in the log form) and an average market capitalization of 4 bn Euro (0.3 bn in the log form). The average board size is equal to 8, both before the announcement of the regulation and also after, suggesting that on average firms do not reduce the number of board seats in order to avoid hiring an extra women and that board sizes are generally sticky. Firm boards have about 14.5% females on average (and 21% post-compliance, unreported), compared to about 6% before the announcement. The 21% average post-compliance is somewhat smaller than any of the quotas considered, since not all regulation is mandatory, and not for all firms in the sample. The main instrument (predicted minimum share of women, $MinShare_i$) averages to 36% and summarizes the average quota-implied share of women in the discontinuity sample. As expected, about half of the firms are located to the right of the kink.

Following prior research on firm value and governance, we compute Tobin's Q as our main measure of firm value (Yermack, 1996; Adams and Ferreira, 2009; Ahern and Dittmar, 2012). It is defined as the sum of total assets and market equity less common book equity divided by total assets, and averages to 1.8 throught the sample. About 19% of firms' capital comes from debt (as normalized to assets), and 71% of firms pay dividends. Average return on assets is slightly negative and amounts to -1%. On average companies' revenue is 0.23 bn per year, and they hire about 1350 workers, for an average labor productivity (revenue per worker) of 238 thousand Euro per worker, with an average wage of about 30 thousand Euro per worker (all values computed based on the log-form averages). There are slightly fewer observations available for other indicators. For the UK firms we also compute loadings on the 4 risk factors, as provided by Gregory, Tharyan, and Christidis (2013), as well as buy-and-hold returns.

Finally, the average age of a director in the sample is 57 years, s/he has on average 1.6 qualifications, a network size of about nine hundred people, and has served in the company for around 7.5 years; about half of directors are independent.

4 First-Stage Results

4.1 The effect of the instrument on the actual share of women

The first empirical test of interest is the one that shows that the instruments (being to the right of the discontinuity, $Right_i$, or the predicted minimum share of women, $MinShare_i$) have a significant and direct impact on the actual share of women, $Share_{it}$. This is a necessary condition for further exploration of the effects of women on corporate outcomes in the IV framework. As discussed above, if firms on average adjust in a different way than substituting women for men, the first-stage coefficient would be close to zero (or even negative). This ultimately becomes an empirical question, which we now explore. To summarize, we find that the instrument does predict differences in female shares, and with a positive sign, implying that on the margin, firms adjust as prescribed by the instrument. In particular, they do not on average choose to change their board sizes to comply with the quota exactly, suggesting that the costs of adjusting board sizes are large enough.

We estimate (1) and (2) and report the results in Table 3. For illustrative purpose only, in columns 1 to 3 we also consider post-announcement vs pre-announcement periods (when dummy $Post_{ct}$ takes the value of 1 for the years after the announcement, and 0 for the years before the announcement), while columns 4 to 6 are estimated on our main post-compliance vs pre-announcement period. Panel A shows the results for the United Kingdom, which constitutes the majority of our observations, and Panel B tracks all countries together. Column 1 uses the simplest possible setup and estimates (1) using the data from the first kink only (i.e. boards of 4 and 5 in the UK). The largest jump in the minimum share occurs at this kink, so the effect on $Share_{it}$ is expected to be the largest. The coefficient of 0.02 implies that there is a 2 pp difference in the share of women on average in the years post-announcement, compared to pre-announcement, between firms that used to have 4 and 5 board members before announcement. We see that while this effect is significant at the 5% level, the instrument is not exceptionally strong (with an F-statistic of 4.5), suggesting that compliance doesn't fully pick up in the first years after the announcement. We therefore move to the post-compliance period, as mandated by the regulation, where all firms likely have had enough time to follow the regulation. As we see in column 4, the respective difference is already 0.065 and significant at the 1% level. If all firms complied exactly with this voluntary regulation in all years and did not change their board sizes, then this magnitude would be 0.15 (the difference between 40% and 25%). However, as noticed before, none of this is assumed in the identification, and it is only important that this instrument does provide a sufficient explanatory power. The economic magnitude of this coefficient suggests that firms comply to a large extent even with the voluntary quota in the UK (and among firms in the FTSE100 compliance rates are the highest at more than 60%). Column 5 repeats the same exercise for firms in all kinks and the magnitude of the coefficient expectedly drops, since the jumps become smaller and smaller, while the instrument is still significant at the 1% level.

In column 6, we turn to using intensities as in (2), where we can fully account for the relative size of the kinks, and this is where the most interesting observation on the economic magnitude comes from. One can think of it as a weighted average of how well firms comply with the instrument. If each firm would satisfy just the minimum required share exactly, as prescribed by the instrument, then the coefficient would have been exactly 1. However, arguably, some firms would prefer to change the board size in the opposite direction (driving the magnitude closer to zero, as discussed above), some would not comply because they are not required to (again, making the coefficient closer to zero), and some may react more strongly and elect a higher percentage than the minimum predicted by the instrument (increasing the magnitude). As such, the obtained coefficients represent a weighted average of all these types of behavior.

In Panel B we consider all firms in our sample together (appropriately accounting for all fixed effects that are now country-specific, as explained in Section 2). As the dynamics of compliance (including the time between announcement and compliance) and the stickiness of boards are likely to be very different across countries, the economic magnitudes may naturally change, but they don't, and the coefficients remain very significant, all at least at the 1% level. The first-stage F-statistic becomes larger than in Panel A in all post-compliance specifications. It is notable that after accounting for all firm heterogeneity and industry-year fixed effects, $MinShare_i$ can still predict the actual share of women quite precisely even across all countries. This is notable because higher quota percentages in other countries also imply that the kinks are located much closer to each other (e.g. 5 vs 6, and 7 vs 8 in case of a 40% quota), and as such there is much less variation left when these firms are compared to each other within such narrowly-defined kinks. Still, we see that our instrument predicts the share of women very well.

4.2 Validating the instrument

4.2.1 Pre-existing differences and dynamics of the first stage

We need to make sure that our instrument is not picking up some pre-existing trends across firms that may relate to future shares of women and future outcomes. We start by exploring visually the dynamics of the first stage in Figures 3 and 4. We plot coefficients from a regression similar to (1) and (2), where instead of $Post_{ct}$ we use the full set of dummy variables for years D_j (the year before announcement, D_0, is excluded to avoid perfect multicollinearity and all coefficient magnitudes are measured relative to this year). Specifically, we estimate:

$$Share_{it} = \gamma_{-4} D_{-4} Right_i + ... + \gamma_{-1} D_{-1} Right_i + \gamma_1 D_1 Right_i + ... + \gamma_8 D_8 Right_i + \lambda_{kt} + \lambda_{st} + \lambda_i + \omega_{it}, \quad (3)$$

$$\begin{aligned} Share_{it} = {} & \gamma_{-4} D_{-4} MinShare_i + ... + \gamma_{-1} D_{-1} MinShare_i + \\ & + \gamma_1 D_1 MinShare_i + ... + \gamma_8 D_8 MinShare_i + \\ & + \lambda_{kt} + \lambda_{st} + \lambda_i + \nu_{it} \end{aligned} \quad (4)$$

and plot the estimates of γ_j with their 95% confidence intervals over time. Since the period between announcement and compliance years varies significantly by country, for illustrative purposes we plot the dynamics for the UK only and highlight 2011 and 2015 on the graph as the announcement and compliance years, respectively. As we observe in Figure 3, the relative difference between firms with closely-held board sizes, γ_j, is statistically the same in each year before announcement as it is in 2010, meaning pre trends are the same for frms on the right and left. This difference also rises steadily starting with the announcement and gets significantly pronounced after the compliance year, validating our primary focus on the post-compliance period. The dynamics for the $MinShare_i$ instrument are depicted in Figure 4 and are also similar.[14]

4.2.2 Pre-existing trends

We now formalize the placebo pre-trend regressions and use all countries. Specifically, we consider the differenced form of (1) and (2), which allows us to use as many years prior to quota announcement for each country as are available in the data (capped at up to 10 years before announcement), and compare the average long-run trends between firms to the left and right of the kink, and firms with different values of the $MinShare_i$ instrument. We estimate:

$$\Delta Share_{it} = \gamma Right_i + \lambda_{kc} + \omega_{it}, \tag{5}$$

$$\Delta Share_{it} = \gamma MinShare_i + \lambda_{kc} + \nu_{it}. \tag{6}$$

The results are reported in Appendix Table 1 for all countries together in columns 1 to 3 and for the UK in columns 4 to 6. As expected, we see no significant effects in any of the specifications, suggesting that there are no pre-existing differences in trends between firms to the left and right of the kink.

We also replicate these regressions for our main second-stage dependent variable – Tobin's Q in Appendix Table 2, columns 1 to 3 for all countries together and 4 to 6 for the UK. Again, we see no significant differences in past trends between firms to the left and to the right of the kink. While this is reassuring and suggests that there is no apparent pre-selection of firms into boards of different sizes, this is also somewhat expected from the way the instrument and the discontinuity sample are constructed to start with.

4.2.3 Why instruments based on past female share cannot be applied

[14] The graphs from similar cross-sectional specifications (w/o firm fixed effects) are reported as Appendix Figures 1 and 2. They are virtually the same meaning that not only past trends, but also past *levels* of the average original share of women are the same for firms on the right and left, and for firms with different values of $MinShare_i$.

To finalize this section, we want to explore why instruments based on pre-announcement shares of women should not be applied in a difference-in-differences setting, at least in our sample. To do that we also consider past trends in Tobin's Q for firms with different pre-announcement shares of women. For illustrative purposes, we first divide firms into those which have at least one female in the year before announcement ($Woman_i = 1$, these are approximately 43% of all firms, meaning most variation is in fact at this threshold) and those which have no women before the quota is announced ($Woman_i = 0$, the remaining 57% of all firms). We then use the share of women as of pre-announcement year, $Share_i$, directly.

As we see in Appendix Table 2, in all countries together (column 7) and in the UK (column 9), firms with at least one woman in the year before the announcement had statistically significantly grown *faster* in terms of Tobin's Q already before the quota was announced, relative to those that had no women. This is also true for the share of women as of pre-announcement year, $Share_i$ (columns 8 and 10). This means that a difference-in-differences setting that is based on assuming such firms would have had the same trends had the reform not happened, will likely find its identification assumptions not satisfied. Specifically, if such an instrument were used to evaluate the effects of the quota, and these differential trends continued to follow, then the reduced form of Tobin's Q on $Post_{ct} * Woman_i$ (or $Post_{ct} * Share_i$) will produce an *upward-biased* coefficient.

The difference-in-differences coefficient of the first stage ($Share_{it}$ on $Post_{ct} * Woman_i$ or $Share_{it}$ on $Post_{ct} * Share_i$), on the other hand, is by construction negative with these instruments. This happens almost mechanically since to get to the same quota firms with more women need to increase their share by less than firms with fewer women. We check that it is also true empirically in our sample. Taken together, this means that the IV coefficient – the ratio of an upward-biased reduced form to a negative first stage – will be *downward biased*: more negative if the reduced form is positive, or less positive if the reduced form is negative. Therefore, using an instrument that uses past shares of women (or past presence of women) as part of its identifying variation will produce a downward-biased estimate, and hence an overly negative view of the effect of women directors on Tobin's Q.

5 The Effect of Gender Diversity on Firm Value

5.1 The effect of Gender Diversity on Tobin's Q

5.1.1 Average effect

We now employ our strategy to estimate the effects of gender diversity on firm performance and other variables. We start by considering Tobin's Q – the most common long-run measure of firm value – as the dependent variable and report reduced-form results (and IV-2SLS) in Table 4.

The reduced form corresponds to the following equations:

$$Y_{it} = \gamma Post_{ct} Right_i + \lambda_{kct} + \lambda_{sct} + \lambda_i + \nu_{it} \tag{7}$$

$$Y_{it} = \gamma Post_{ct} MinShare_i + \lambda_{kct} + \lambda_{sct} + \lambda_i + \nu_{it}, \tag{8}$$

and the IV-2SLS is given by:

$$Y_{it} = \beta Share_{it} + \lambda_{kct} + \lambda_{sct} + \lambda_i + \kappa_{it}, \tag{9}$$

where the instrument is either $Post_{ct}Right_i$ (columns 1 and 2 – for all countries, and 4 and 5 – for the UK) or $Post_{ct}MinShare_i$ (column 3 – for all countries, and 6 – for the UK), and all the variables and fixed effects are defined as above.

We include industry-country-year fixed effects in all specifications (based on Eikon 52 industry groups), to make sure the differences in Q are not accidentally driven by non-random composition of board sizes across different industries and shocks to them, as well as to explain more variation in Tobin's Q. The coefficient 1.011 in column 1 suggests that firms to the right of the threshold (at the first kink), i.e. those with one more woman, have on average 1.011 units higher Tobin's Q than those to the left of the discontinuity, after quotas were introduced vs any potential difference before. In column 2 we replicate this analysis across all kinks and see similar results. In column 3, we move to the second instrument, $MinShare_i$, which is based on intensities, and again we see very significant reduced-form results.

By assuming the exclusion restriction that being on the right or left of the threshold affects firm outcomes only through the different share of women to have, we can additionally convert our estimates into IV-2SLS and obtain the magnitude of the effect, rather than just its sign. We discuss the validity of exclusion restriction further in Section 6.1.2.

Just below, we also report the corresponding IV-2SLS coefficients with their standard errors, as well as robust weak-IV Anderson-Rubin confidence sets, which provide the accurate p-values when F-statistics are low (Andrews et al., 2019). While the reduced-form coefficients are obviously different in magnitude across specifications, once we rescale them into the actual magnitudes of interest – the IV-2SLS effects of women on firm value, – we see very similar magnitudes across all three instruments. This is quite remarkable, given that they are based on slightly different samples (largest kink vs all kinks) and slightly different identification assumptions, suggesting that this average positive effect of women on Tobin's Q is very robust. The preferred IV-2SLS estimate of 20.76 in column 3 means that for a one-half of a within-firm standard deviation in the share of women (i.e. by 0.05), Tobin's Q on average increases by 1 unit (which is about 0.6 within-firm standard deviations of this variable).[15] This suggests that women do have an economically large effect on Tobin's Q, across European listed firms. The Anderson-Rubin p-values further confirm that this effect is significant at the 5% level in each of the three specifications. In columns 4 to 6 we redo the analysis for the UK only, and the results are similar.

These magnitudes, while appearing large, are sensible for at least three reasons. First, they are local in the sense of being identified from small changes in the instrument, $MinShare_i$ (which at most varies in the 25-40% range) and the corresponding changes in the actual share of women, $Share_{it}$ (which is about 6.5% for crossing the largest kink, as per column 4 Table 3). As such, these magnitudes cannot even in principle be extended to changes from e.g. an all-male to an all-female board, unless exceptionally strong linear extrapolations are made. Our IV-2SLS estimates thus should only be interpreted with respect to smaller changes in the share of women. Second, for all

[15] Given that all of our regressions are panel within-firm, we think within-firm standard deviations are a more relevant benchmark for interpretations. However, using overall standard deviations (as of Table 2) provides a similar answer: an increase of $Share_{it}$ by one-half of an overall SD (i.e. by 0.0775) leads firms to increase their Tobin's Q by 1.6 (which is about 0.55 of the overall SD of this variable).

of our kinks, being on the right also corresponds to having at least 3 women on the board (or at least 35% of women for smaller boards), which is interpreted as achieving a critical mass that is necessary for becoming more active boards (Schwartz-Ziv, 2017).

Third and finally, our effects can be directly discussed in the framework of heterogeneous treatment effects, where IV estimates the local average treatment effect on compliers (Imbens and Angrist, 1994). Within this framework, our IV estimates correspond to firms that "comply" with our instrument, i.e. those that have one more women *only* because they end up to the right of the threshold with the rounding not in their favor, and they would not have had this extra woman, had they ended up to the left of the threshold.[16] These are likely to be firms with the stickiest board sizes, i.e. firms for which the costs of having more women are smaller (or non-existent) relative to the costs of changing the board size. It is thus expected that this type of firms may have quite high positive effects of women directors. While it is inherently impossible to observe which firms are compliers to the instrument, we can calculate how many they are in our sample. To do that, in unreported results, we estimate (7) with the dummy of having the share of women at or above the quota as the dependent variable.[17] We observe that the percent of instrument compliers varies from 13% across all countries and all kinks to 18% in the UK for firms at the first kink. It is reassuring that despite a seemingly small variation in our instrument, this percentage of compliers is rather significant. It is also similar to the percent of instrument compliers (15% across all countries) that would be achieved if one used the $Post_{ct} * Woman_i$ instrument, discussed above, even on *all* board sizes. This means that instruments based on pre-existing heterogeneity of female shares are likely to neither produce consistent estimates, as discussed in Section 4.2.3, nor apply to larger subsamples of firms.

5.1.2 Individual country-level analysis

We now further decompose the overall average effect from Table 4 to see if it is driven entirely

[16] All previous IV-based work in this literature is of course also subject to such a LATE interpretation. These compliers to the instrument are of course not to be confused with voluntary compliers to the regulation itself.

[17] This is a virtue of having binary variables: in the Angrist and Pischke (2009) notation, this is a regression of the endogeneous treatment indicator on the instrument that automatically calculates the proportion of instrument compliers. See there also for the definition of instrument compliers, which are defined as of counterfactuals, and are a different concept from the actual observable compliance to the percentage regulation.

by the UK or if there is evidence from other countries as well. The very basic obstacle to analyzing country-by-country is the individual-country-level first stage, as there are multiple reasons why it may have different strengths across countries, especially given a small number of listed companies in most countries. First, the regulation is different across countries in terms of how obligatory compliance is de jure and de facto, i.e. whether the sanctions for non-compliance are significant enough to alter firms' behavior in the institutional environment of a specific country. Second, the speed of compliance with the quota may be different, depending on how easy it is to change board members, and how big the lag is between announcement and actual compliance date, both affecting the timing of the first-stage effects. Finally, as discussed in Section 2, the exact way how firms adjust to the quota (by substituting women for men or hiring additional women until the quota is satisfied or a combination of the two) directly affects the value of the first-stage coefficient. Given these many obstacles, we can only explore the effect of interest in countries or subsamples of countries where the first stage proves significant enough.[18]

To increase the power of the first stage at the individual-country level we estimate our main equations of interest using post-compliance vs pre-announcement periods, for the first two kinks.[19] We report the results in Table 5, where Panels A and B correspond to the $Right_i$ and $MinShare_i$ instruments, respectively. The coefficients within each column refer to the results of separate regressions: the first stage in the first row, the reduced form in the second row, and below we also report the the implied 2SLS estimate as well as the Anderson-Rubin robust confidence sets with corresponding p-values. In column 1 we consider all countries together, and in column 2 – just the UK. We provide these for reference only to check that our main results from Tables 3 and 4 are not affected by this change towards a more powerful sample.

Now in column 3 we report the effect on all countries except the UK. As we see in Panel A,

[18]One may be tempted to simply explore the effect of the instrument on Tobin's Q (the reduced form), even if that first stage is not powerful enough. However, this is not reasonable in our setting, because even the sign of the first stage is unknown (as it depends on the way firms adjust to the quota). Hence there is no reason to expect any specific sign from the reduced form either.

[19]For the individual country-level analysis, it becomes even more important to keep a sufficient number of observations to have a high enough power to detect variation in the female ratio. Two largest kinks is therefore a reasonable balance between using the jumps with the highest magnitude, yet without further restricting the number of observations to just the first kink. Nevertheless, we check that our results (unreported) are similar when using only the first kink: for all countries where the first stage is positive and significant at 5% level, IV-2SLS coefficients are also positive and significant at at least 10% level.

both the first stage (0.640) and the reduced form (0.391) are significant at the 5% level and have the expected signs. However, the first-stage F-statistic of 6.60 is not high enough for 2SLS to provide credible inference, and the magnitudes of IV-2SLS cannot be directly compared (with weak instruments 2SLS is biased towards OLS which is close to zero in our case). Therefore, we have to refer to Anderson-Rubin confidence sets and p-values that are robust to the presence of arbitrarily weak instruments (Andrews et al., 2019), to see if there is a significant effect. As we see, the 95% confidence set does not include zero, meaning that the effect of female share on Tobin's Q is also positive and significant at the 5% level for all countries other than the UK, with at p-value of 2%. This effect is also robust to using the $MinShare_i$ instrument instead (Panel B), with a corresponding p-value of 1.2%. Overall, we see that the effect of female share on Tobin's Q is pronounced both in the UK and in all non-UK countries taken together.

Furthermore, there are three more individual countries for which the first stage coefficient turns out to be significant at least at the 5% level, and hence we can explore the second-stage effects for them: these are France, the Netherlands, and Norway (columns 4 to 6, respectively).[20] As we see, all three show positive effects in the reduced form, with France and Norway being significant at the 5% level, while for the Netherlands the coefficient has a 10.5% p-value. A similar pattern is observed once we consider AR confidence sets. Both France and Norway show significant positive effects of female directors on Tobin's Q, while the effect is positive but only marginally significant for the Netherlands (p-value of 10.6%). For the Netherlands and Norway, which have a relatively high F-statistic, we can look directly at the implied IV-2SLS estimates, and it is interesting to see Norway having a very similar magnitude to the one in the UK. For the Netherlands the magnitude is lower, but it is clearly not negative (rejected at the 10.6/2=5.3% significance level). The effect for France is also somewhat smaller in magnitude than that for the UK. This may be explained by quotas being imposed on non-executives only there, or by other institutional differences across countries, as mentioned above. Nevertheless, since AR confidence sets intersect for any two sets of countries, the differences across them are not statistically significant. Hence we prefer not to speculate further on cross-country heterogeneity. Finally, in column 7 we report the results for

[20]For each other country considered individually, the first stage is not significant even at 5%. Hence the second stage cannot be identified, as the AR confidence sets would automatically include plus and minus infinity.

all other countries together, for completeness only: the first-stage F-statistic of 0.01 prevents any separate inference for these countries, even though it is reassuring that the reduced-form coefficient of 0.0464 is also positive. Panel B replicates the results using the $MinShare_i$ instrument, which may capture relative intensities better, and the results are robust.[21]

To sum up, we find no evidence of women affecting firms' value negatively. If any, the evidence for both the UK and all other non-UK countries together demonstrates a significant positive effect on Tobin's Q. Finally, the results for individual countries where the effect can be in principle explored also show a positive effect.

5.2 Is this really about firm value?

While there is a strong and robust positive effect of female directors on Tobin's Q, this measure has been somewhat criticized for not being the best one (see e.g. Dybvig and Warachka, 2015). Since we can use the same shock to female directors to study any measure of performance, we also look at the ultimate measure that investors would earn – buy-and-hold returns (Erkens et al., 2012) – in Section 5.2.1. Then we explore the dynamics of annual abnormal returns – in Section 5.2.2. Finally, we investigate if these effects are accompanied by a different risk-taking profile of firms with more and fewer women – in Section 5.2.3. In this section, however, we have to refer to the UK only, which is where we are able to locate country-specific risk factor returns, that are needed for parts of our analysis.

5.2.1 Long-run buy-and-hold returns

Figures 5 and 6 plot the buy-and-hold returns that investors would earn by each date if they bought a portfolio of firms to the right (solid line) or to the left (dashed line) of the kink a year before the announcement of the regulation (February 24th, 2010), for the first and all kinks, respectively. As we see, if investors bought and held firms to the right, they would have earned a much higher return over these years than had they bought and held firms to the left. Specifically, the difference

[21] These cross-country results also demonstrate that our main results cannot be explained by different trends in odd-number-sized boards which may have a higher effectiveness of per se (see e.g. Deng et al, 2012), since whether an odd-numbered board ends on the right or left depends on the quota percentage (e.g. right for the UK and Netherlands, left for France, and mixed for Norway).

amounts to about 3.8% in the average annual compounded return for firms at the first kink, and for 1.6% across all right and left firms.[22] This means that investors win in the long-run from having more women in the board.

We further support this conclusion in Table 6 by regressing these buy-and-hold returns on our instrument, $Right_i$, with a separate coefficient in each year. To do that, we estimate the following equation:

$$BH_{it} = \gamma_{2011} Year_{2011} Right_i + ... + \gamma_{2019} Year_{2019} Right_i + \lambda_{kt} + \lambda_{st} + \lambda_i + \nu_{it}, \qquad (10)$$

where BH_{it} is the buy-and-hold return of firm i in year t (i.e. the total return that an investor would earn if she held this stock till year t), $Right_i$ is defined as before, and $Year_j$ are the indicator variables for each particular year j after announcement. The year before announcement (fiscal from February 24th, 2010) is the base category when all $Year_j$ are zero and buy-and-hold returns of all firms are mechanically set to 1. To make our results more comparable to the previous section, we also add kink-year fixed effects λ_{kt} (to make sure we compare only within closely-held board sizes and not across), firm fixed effects, λ_i, and industry-year fixed effects, λ_{st}, which make sure we compare firms within the same industry, and that our results are not explained by e.g. firms to the right accidentally being located in industries that experienced a boom during this period. Note that all $Year_j$ indicator variables (without the interaction) and $Right_i$ itself are absorbed automatically by λ_{st} and λ_i fixed effects, respectively. This means that the interpretation of each γ_j is the difference in buy-and-hold returns of firms on the right of the kink relative to firms on the left of the kink, in year j. This specification is also more saturated than the simpler graphical evidence in Figures 5 and 6, because now firms are also compared within the same kink and industry, further reducing any concerns for potential omitted factors driving these returns.

Panel A reports the results for firms with 4 and 5 board members, while Panel B considers all firms together. As we observe, the differences in buy-and-hold returns are not only economically large, but also accumulate to a statistically significant difference for most years for 4 vs 5 firms,

[22] The average annual compounded return for firms with 5 board members is 10.12% = 2.38^(1/9)-1 and for firms with 4 board members it is 6.33% = 1.74^(1/9). All firms on the right earned 10.31% = 2.42^(1/9), and all firms on the left earned 8.69% = 2.12^(1/9).

and in later years for all firms, and the magnitudes are almost identical to what we have just seen in the graphs (without controls).

5.2.2 Long-run annual abnormal returns

The comparison of buy-and-hold returns is essentially a cross-sectional one, since by construction the buy-and-hold returns of all firms in the beginning of investment period are set to 1. We therefore want to make sure they are not a result of some pre-existing differences between firms on the left and right. Additionally, buy-and-hold returns cannot precisely illustrate the dynamics of the effect, because they are cumulative, and hence any past stock price surprise affects all future returns, and in a non-linear way. For example, it would be strange if investors kept getting positive surprises from the fact that firms have a higher share of female directors on their boards: such a finding would instead be indicative of potential omitted variable concerns. To refute both of these concerns, we turn to the difference-in-differences specification on annual (non-cumulative) returns, and estimate the following equation:

$$R_{it} = \gamma_{2011} Year_{2011} Right_i + ... + \gamma_{2019} Year_{2019} Right_i + \lambda_{kt} + \lambda_{st} + \lambda_i + \nu_{it}, \tag{11}$$

where R_{it} is the annual return of firm i in year t (i.e. the total return that an investor would earn if she held this stock till year t), and all other variabels are defined as before. We now estimate this specification from three years before annoucement (as in the main analysis of the paper in Section 5.1). This gives each γ_j a proper difference-in-differences interpretation: it is the difference in annual returns of firms on the right of the kink relative to firms on the left of the kink, in year j, *over and above* any such difference in the three years before annoucement. As such, the returns are abnormal relative to the standard pre-existing difference between firms on the right and left. Because we include industry-year and kink-year fixed effects, they are further abnormal to any observed and even *unobserved* risk factors common to the industry, and those common to the kink.[23]

Panel A of Table 7 reports the results for firms with 4 and 5 board members, while Panel B considers all firms together. As we see, the positive stock price surprises happen only in the year

[23] As Gorovyy et al (2021) argue, accounting for risk factors using group-time fixed effects has the advantage of being able to take care of loadings on all observed and even unobserved factors common to the group. In our case, this means that any $\beta_s F_t$ is automatically absorbed by λ_{st}, and any $\beta_k F_t$ is absorbed by λ_{kt}.

of annoucement and (to a smaller and less significant extent) in the year of compliance.This is reassuing since it is the expected dynamics of price adjustment, consistent with investor surprises, rather than omitted factors that may have shown up as abnormal returns every year. In terms of the economic magnitudes, these two spikes of 10-13% in 2011 and 6-8% in 2015, coupled with negligible statistically zero returns in other years, on average yield precisely 1.5-2% higher average (non-compounded) returns over the 9 years of consideration. The magnitudes are thus consistent with those just presented in Section 5.2.1 with slight differences due to non-linear compounding and the move from cross-section to difference-in-differences.

5.2.3 Are the positive value effects accompanied by higher risk taking?

Finally, we want to know if the positive long-run value effects are accompanied by the different risk-taking profiles of firms with more and fewer women, as can be measured by their differential loadings on common risk factors. To do that we calculate betas with respect to the 4 factors for each firm for each calendar year (with a minimum of 100 days of non-missing observations within that year), and estimate the same equation as in (7), but now using these betas as dependent variables.[24] The coefficients thus measure whether firms to the right of the kink changed their betas by more than firms to the left of the kink, post-compliance relative to pre-announcement periods.

The results are reported in Table 8 for firms at the first kink (Panel A) and all firms together (Panel B). As we see, for neither of the 4 risk factors did the loading of firms to the right change differentially over time compared to that of firms to the left. For robustness we also check the post-announcement period relative to pre-announcement (Appendix Table 3). As we see, there is only one significant coefficient at the 10% level (higher momentum loading of firms to the right in the specification of all firms). Given our 16 specifications in these results, this is expected and likely to be an artefact of type I error, especially since this difference is not present for the firms to the right of the first kink, which have the highest return premium. Nevertheless, we do an additional back-of-the-envelope calculation and find that the magnitude of this difference is far enough from explaining the difference in buy-and-hold returns between firms to the right and left.[25] All in all,

[24] We use years up to 2017, as this is the last year when UK factor returns are available. We also do not include industry-year fixed effects in these specifications because of the nature of individual stock beta coefficients, as well as to increase power when observations are relatively few.

[25] According to Gregory et al (2013) Table 9, the price of risk for a unit of momentum factor loading was about 0.58

this means that the higher buy-and-hold returns of firms to the right are not accompanied by higher risk profiles post reform. We will get back to this finding once we discuss the mechanism in Section 6.2.1.

6 The Effect of Gender Diversity on Other Firm and Board Characteristics

So why does the gender reform bring positive value effects? In this section we explore the two facets to it that have been proposed as candidates in the literature. Specifically, in Section 6.1 we look at board characteristics and behavior, while in Section 6.2 we study firm strategies and outcomes.

6.1 Does board composition change?

6.1.1 Average director attributes and attendance

We start by exploring the average characteristics of the board members, such as age, number of qualifications, share of independent non-executive directors, network size, and time in the company, as some of these have been proposed as potential mechanisms (see e.g. Bertrand et al., 2019, Ferreira et al., 2020). The results of estimating (8) with these dependent variables are reported in Table 9, with Panel A for all countries and Panel B for the UK.[26] As we see all results consistently indicate that the average characteristics of the boards with relatively more women (those on the right) are the same as those with fewer women (those on the left).[27] If anything most coefficients are negative (insignificantly), indicating that there are no differences in the attributes of the incoming women directors, and hence they are unlikely to be responsible for the positive value effects.

in the UK (significant at 10%). This means that a 0.06 higher beta implies a 0.035%(= 0.06*0.58) higher monthly return, or about a 0.4% higher annual return. This is much smaller than our effects in Section 5.2.1.

[26] Because the number of observations varies dramatically across various dependent variables, the availability of IV-2SLS estimates depends on the strength of the first stage, which can be low in small samples. We therefore instead opt to consistently present the reduced-form results in this and all the following sections, as they are not dependent on the precision of the first stage. For the same reason, we opt to keep all countries together, rather than separating the non-UK countries from the UK. In specifications with a large number of observations, where the first stage is precisely estimated, the IV-2SLS magnitudes can be obtained as Wald estimates by dividing the reduced-form coefficients by the respective first-stage estimate from Table 3.

[27] To save space, in Tables 9 to 12 we report the results with $MinShare$ instrument only. The results with $Right$ are very similar and available upon request.

We can also measure one behavioral response – average board meetings attendance; we report these results in column 7. First, we see that average board meeting attendance significantly increases with more women, consistent with the evidence in Adams and Ferreira (2009) and board with a critical mass of women being more active (Schwartz-Ziv, 2017). Although the average percentage attendance cannot proxy how efficient these meetings are, we can back up the economic magnitude of this coefficient. Dividing the 18.12 in column 7 by the corresponding first-stage estimate of 0.319, we get a magnitude of 56.8. This means that a 10pp increase in the actual share of women, induced by the instrument, increases attendance by 5.68 pp. Given the overall average of about 94%, this looks quite small. However, it also means that in many firms directors will shift from missing one time out of 20 towards full attendance. This does not look large in real-life terms, though it may be if the full discipline comes only when nothing is ever missed. As such, board meeting attendance appears to be at least part of the mechanism of how having more women on the board increases value.

6.1.2 Board size and the exclusion restriction

As we have seen from the first-stage results, board size appears to be sticky enough to make firms marginally prefer switching men for women. However, some firms may decide to adjust the board size as a response to the regulation: firms to the right of the kink may choose to remove one board member in order to not have to hire too many women directors, making compliance easier. While this additional indirect response to the quota is an interesting and important question itself, one may also have a concern that a smaller board size per se increases performance (see e.g. Jenter, Schmid, and Urban, 2019, for recent quasi-experimental evidence), and as such violate the exclusion restriction of the instrument. We report the results of estimating specifications (7) and (8) for the board size as a dependent variable in Appendix Table 4. We see that there is also some downward adjustment of the board size on average, as indicated by columns 2 and 3, suggesting that regulation additionally induces a change in the optimal board size.

One simplest way to take care of it is to just stick to the reduced-form results, and interpret all of our findings as the effects of gender regulation itself. However, we can go further and instead calculate how much of the IV estimate can be attributable to the violation of the exclusion restriction through board size adjustments. To do that we first need an estimate of the direct effect of board size on performance, such as the one presented by Jenter et al. (2019). Their magnitudes suggest that one additional member of the board increases Tobin's Q by about 0.05-0.06. Multiplying this by the largest possible estimate in Appendix Table 4 (-2.677 in column 3), we get an indirect effect of about 0.13-0.16 on Tobin's Q, had it been explained solely by the board size adjustment channel. Our main estimate of 6.615 in column 3 Table 4 is clearly more than an order of magnitude larger, suggesting that only about one fiftieth to one fortieth of the total positive effect of the reform on firm value can be attributed to board size adjustments. This back-of-the-envelope calculation is also robust to using other estimates from the literature.[28]

Finally, to complement this calculation, we also perform the full mediation analysis by including the contemporaneous board size as a control variable directly. These results are reported in Appendix Table 5 and they are almost identical to those reported in the main Table 4. While the first-stage F-statistics expectedly become a bit lower, the second-stage results are the same, both in magnitude, and in significance (with AR p-values of 1-6% depending on the specification). Interestingly, the coefficient at board size control (0.02-0.13, depending on the specification), albeit not very precisely estimated, is also similar to the magnitudes reported by the previous authors. All in all, these two bits of additional analysis suggest that board size adjustments cannot explain the increases in firm value after the reform.

6.2 Performance decomposition and empire building

As we see, the channels that can impact firm value through board characteristics and behavior seem limited in terms of their economic magnitudes. We therefore turn to firm strategies and outcomes.

In this most common definition, Tobin's Q can be decomposed into $1 + MV/TA - BV/TA$, i.e. the difference between the ratios of market value of equity to total assets and book value of equity

[28] For example, one of the earliest works by Yermack (1996) implies that reducing board size by one member from 5 to 4 (or by 20%) associates with increases in Tobin's Q by 0.2*0.337=0.0674, which is very close to Jenter et al. (2019) experimental estimates.

to total assets (plus 1). To see which of the two parts drives the main result, we estimate our main specification 8 with these two ratios as dependent variables in Table 10 for all countries (Panel A) and the UK separately (Panel B). We observe that it is the first part that mostly contributes to an increase in Tobin's Q: market value of equity to assets rises by about 5.3, while book value of equity to assets drops by about 0.3 (insignificantly), for every 10pp of the expected minimum share of women on boards (or 3pp in the actual share of women). These coefficients are similar across panels.

So why do firms have higher market, but similar or lower book values at the same time? This can be consistent with at least four (non-mutually exclusive) explanations: higher leverage, higher dividends paid, decrease of scale (e.g. writing off some unproductive assets), and a temporary negative performance shock (that drills down retained earnings). In Table 10 we consider each of them. We observe that firms do not increase debt-to-assets ratio (column 3). Nor do they increase dividends: neither in the dividend yield (column 4), nor in dividend payout or propensity of paying any dividend (unreported for brevity). If any, the evidence suggests the opposite: firms with more women pay fewer dividends. This suggests that the first two explanations are not responsible for the observed effects on Tobin's Q.

At the same time, there is strong evidence that both assets and operating return on assets fall (columns 5 and 6). This means that there is a negative effect on the size of the firm and its operating performance, and it is therefore worth exploring both the asset and operating sides in more detail.

6.2.1 Are women less prone to empire building?

The literature has shown that men are more overconfident (see Croson and Gneezy, 2009, for a survey) and that overconfidence itself leads to more investment distortions and empire building (Malmendier and Tate, 2005). This means that the potential positive effect of more women on firm value may in fact come from less empire building activity (see e.g. Levi et al., 2014, for documenting a correlation between gender and M&A activity). We now explore this channel using our causal framework.

In Table 10 we observe that total assets rise by significantly less in firms with more women.[29]

[29] Note that here and in what follows we are being careful to interpret negative coefficients at all variables that

Specificaly, they do so by about 20% less for every 10pp of the expected minimum share of women in boards (or 3pp in the actual share of women). To see if this is explained by less empire building, we look at several measures of investment activities. Specifically, in Table 11 we examine indicator variables for whether the firm spent money on acquiring a business (merger-related expenses), whether it received money from selling a business (including discontinued business units, branches, and divestitures), whether it purchased fixed assets, and whether it received a cash inflow from selling fixed assets.

As we see in column 1, firms with more women on their board are relatively less likely to acquire a business: for every 10pp expected minimum share of women (or 3pp in the actual share of women), firms are 10pp less likely to have incurred any merger-related expenses, which is a large economic magnitude (yet, as before, we warn against linear extrapolations to larger changes in female share).[30] Interestingly, this effect is not about switching the types of business, since there is no simultaneous change in sales of business (column 2). We observe very similar dynamics in the fixed assets: firms with more women are less likely to purchase new fixed assets (column 3), but no less likely to generate cash from selling them (column 4). This asymmetry is interesting in that it suggests that firms with more women are not just writing off some unproductive assets, but instead they are less likely to buy these assets to start with. There are two potential explanations consistent with this evidence: a reduction in overall empire-building activity and a reduction in diversifying acquisitions (as in Gormley and Matsa, 2016). However, given our evidence in Section 5.2.3 on no increases in risk, and on positive market reaction in general, we are left with the former explanation of women being less prone to empire building.

6.2.2 Operating performance

Since the return on assets is much lower for firms with more women (column 6 of Table 10), we further investigate the operating performance side and present the results in Table 12. First, we observe that sales (both as a logarithm, in column 1, and as a share of assets, in column 2) rise by significantly less, with a magnitude of about 30% for each 10 pp increase in the predicted

increase over time as a "lower increase" rather than a "decrease".

[30] As of standard deviations, an increase of $Share_{it}$ by one-half of an overall SD (i.e. by 0.0775) leads firms to decrease their acquisition propensity by 0.0805 (all countries), which is about 0.2 of the overall SD of this variable.

share of women (or 3pp in the actual share of women). At the same time, there is no significant increase in the operating expenses to assets ratio (in column 3), nor is there any increase in R&D or labor expenses, as a share of assets (unreported for brevity, yet the number of observations is much smaller for these variables). This means that the change in OROA is mostly driven by a lower increase in sales, rather than worsening of the costs side. This is further supported by the observation, that none of the profit margins fall: both gross and operating profit margins (as a share of sales) are stable (columns 4 and 5).

As such, while lower OROA is consistent with the evidence in Matsa and Miller (2013), the underlying reason is different (lower increase in sales rather than higher labor expenses). To further observe that, in column 6 we see that employment also rises by significantly less, by about 20% for each 10 pp increase in the predicted share of women (or 3pp in the actual share of women), but not the labor productivity (column 7) or average wage (column 8). This means that the quality of workers does not deteriorate. All these results are important together, since they indicate that the primary driver of the fall in OROA are the lower increases in sales (which are accompanied by all other proportional changes). At the same time, none of the sales-based margins fall, suggesting that the company makes the same profit out of each unit sold and that workers are similarly productive and paid similarly.

As Matsa and Miller (2013) discuss, such differences in the return on assets would likely diminish over time. And indeed, by examining a longer sample, Eckbo et al .(2021) show that the negative ROA effects disappear over time. Additionally, any perturbation of the board per se, even the one unrelated to gender, is likely to affect performance in the short run (see e.g. Nguyen and Nielsen, 2010, on the negative reaction to sudden director deaths). Thus, the remaining question is whether these changes are in fact short-run or not. Most countries have only recently implemented the regulation, so we haven't yet lived to observe and analyze the more long-run data. What we can do now is to explore the yearly dynamics, and specifically for the most interesting variable, which is the main driver of the operating performance decreases, – sales. We therefore plot the dynamics for the logarithm of sales, based on a yearly regression similar to (3), in Figure 9 (we do so for the UK, again because different countries have different durations between announcement and compliance).

As we notice, there are some slight upward dynamics in the later years, and sales in the later two years are not significantly different from the whole pre-announcement period. This is suggestive of the shock to sales being more transitory, rather than permanent.

To sum up, our evidence is consistent with firms with more women on the board being less prone to empire building. The smaller increases in assets in firms with more women are accompanied by more than proportionally smaller increases in sales that drive some of the operating performance indicators, such as OROA, down. However, all profit margins are stable, and workers do not become any less productive, suggesting that it is indeed the relatively smaller increase in sales that drives everything down. This shock is likely to be temporary, and later research will be able to explore that in more detail. Importantly, none of these changes are accompanied by negative market reactions. If anything it is positive, suggesting that these operating changes are viewed positively by the market, consistent with firms with more women engaging less in scaling up the inefficient operations.

7 Conclusion

In this paper we explore the effects of increased female presence on corporate boards on value, operating performance, and other firm and board characteristics, for a set of European countries that introduced soft or hard regulation with respect to the share of women. While previous research has extensively looked at these questions using instruments based on past female shares, we show that they cannot be applied, at least in our sample of all countries, because firms with women and those without already grow at different rates before any regulation is introduced. This difference in trends would thus produce an overly pessimistic view of the effects of women on corporate performance.

Instead, we use a novel identification strategy that allows us to estimate causal effects under the minimum possible assumptions in a setting with a universal quota or recommendation. We find evidence of positive effects on the value of the firm (as measured by Tobin's Q and buy-and-hold returns). We further dig into possible mechanisms and observe that these positive effects are not explained by higher riskiness of firms or higher-quality boards (except for a slight positive effect on board meetings attendance). The main driver of our effects is the reduction in empire building, as proxied by lower inreases in assets, lower merger-related expenses and less investment in purchases of fixed assets. These effects are accompanied by a disruption in sales, which drives some of the operating performance indicators down, likely temporarily. However, since margins and labor productivity are stable, and these operating performance decreases are not accompanied by lower market values, this is further evidence that firms with more women engage less in inefficient operations.

Our results have important policy implications. With a general socially-based move towards gender equality, many countries have pushed quotas in corporate boards, yet the effects on shareholder value and firm policies have been debatable. We show that there is no negative effect on value, and boards do not become any less competent with more women on boards. Having more women on boards, besides being socially important, is also aligned with corporate interests.

References

[1] Adams, R. B., and D. Ferreira, 2009, "Women in the Boardroom and Their Impact on Governance and Performance," *Journal of Financial Economics*,94(2): 291–309.

[2] Adams, R. B., and P. Funk, 2012, "Beyond the Glass Ceiling: Does Gender Matter?" *Management Science* 58 (2): 219–235.

[3] Adams, R. B., B. E. Hermalin, and M. S. Weisbach, 2010, "The role of boards of directors in corporate governance:A conceptual framework and survey, " *Journal of Economic Literature* 48: 58–107.

[4] Ahern, K. R., and A. K. Dittmar, 2012, "The changing of the boards: The impact on firm valuation of mandated female board representation,". *Quarterly Journal of Economics* 127:137–197.

[5] Andrews, L., J. H. Stock, and L. Sun, 2019, "Weak instruments in instrumental variables regressions: Theory and practice," *Annual Review of Economics* 11, 727-753.

[6] Angrist, J. D., G. W. Imbens, and D. B. Rubin, 1996, "Identification of causal effects using instrumental variables," *Journal of the American Statistical Association* 91 (434): 444-455.

[7] Angrist, J.D. and V. Lavy, 1999, "Using Maimonides' rule to estimate the effect of class size on scholastic achievement," *Quarterly Journal of Economics* 114(2), 533–575.

[8] Angrist, J.D. and J-S. Pischke, 2009, *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press

[9] Bertrand M., S. E Black, S. Jensen, A. Lleras-Muney, 2019, "Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labour Market Outcomes in Norway," *Review of Economic Studies* 86(1), 191–239.

[10] Co-determination Act of May 4, 1976 (Federal Law Gazette I p. 1153), last amended by Article 7 of the Law of April 24, 2015 (Federal Law Gazette I p. 642)

[11] Carhart, M. M., 1997, "On Persistence in Mutual Fund Performance," *Journal of Finance* 52, 57–82.

[12] Carter, D. A., B. J. Simkins, and W. G. Simpson, 2003, "Corporate governance, board diversity, and firm value," *Financial Review* 38 (3): 33–53.

[13] Coles, J. L., D.D. Naveen, and L. Naveen, 2008, "Boards: Does One Size Fit All?," *Journal of Financial Economics* 87 (2), 329–356.

[14] Comi S., M. Grasseni, F. Origo, L. Pagani, 2020, "Where Women Make a Difference: Gender Quotas and Firms' Performance in Three European Countries", *ILR Review* 73 (3).

[15] Croson, R. and U. Gneezy, 2009, "Gender differences in preferences," Journal of Economin. Literature 47, 448–474.

[16] Lord Davies of Abersoch, CBE, 2011, *Women in boardrooms*, February 2011

[17] Deng, X., H. Gao, W. Liu, 2012, "Voting Efficiency and the Even-Odd Effects of Corporate Boards: Theory and Evidence", *working paper.*

[18] Dybvig, P. H., and M. Warachka, 2015, "Tobin's Q Does Not Measure Firm Performance: Theory, Empirics, and Alternatives," *working paper.*

[19] Eckbo, B. E., K. Nygaard, and K. S. Thorburn, 2021, "Valuation effects of Norway's board gender-quota law revisited," *Management Science.*

[20] Erkens, D.H., M. Hung, and P.Matos, 2012, "Corporate governance in the 2007–2008 financial crisis: Evidence from financial institutions worldwide", *Journal of Corporate Finance* 18(2), 389–411.

[21] European Comission, 2016, *Gender balance on corporate boards: Europe is cracking the glass ceiling*, July 2016 Fact sheet

[22] Ferrari G., V. Ferraro, P. Profeta, C. Pronzato, 2016, "Gender Quotas: Challenging the Boards, Performance, and the Stock Market," *IZA DP 10239.*

[23] Ferreira, D., 2010, "Board Diversity," in *Corporate Governance:A Synthesis of Theory, Research, and Practice* (eds H.K. Baker and R. Anderson).

[24] Ferreira, D., 2014, "Board Diversity: Should We Trust Research to Inform Policy?" *Corporate Governance: An International Review* 23(2):108–111.

[25] Ferreira, D., E. Ginglinger, M.-A. Laguna, and Y. Skalli, 2020, "Board Quotas and Director-Firm Matching," *working paper.*

[26] Gertsberg M., Mollerstrom J., Pagel M., 2021, "Gender quotas and support for women in board elections," *working paper.*

[27] Gormley, T. A., V. K. Gupta, D. A. Matsa, S. C. Mortal, and L. Yange, 2020, "The Big Three and Board Gender Diversity: The Effectiveness of Shareholder Voice," *working paper.*

[28] Gormley, T.A., and D.A. Matsa, 2016, "Playing it Safe? Managerial Preferences, Risk, and Agency Conflicts," *Journal of Financial Economics* 122, (3): 431-455.

[29] Gorovyy S., P. Kelly, and O.Kuzmina, 2021, "Does Secrecy Signal Skill? Own-Investor Secrecy and Hedge Fund Performance," *Journal of Banking and Finance* 133: 106288.

[30] Gregory A., R. Tharyan, and A. Christidis, 2013, "Constructing and Testing Alternative Versions of the Fama–French and Carhart Models in the UK," *Journal of Business Finance and Accounting* 40 (1) and (2), 172–214.

[31] Greene D., V. J. Intintoli, K. M. Kahle, 2020, "Do board gender quotas affect firm value? Evidence from California Senate Bill No. 826," *Journal of Corporate Finance* 60.

[32] Hoogendoorn S., H. Oosterbeek, and M. van Praag, 2013, "The Impact of Gender Diversity on the Performance of Business Teams: Evidence from a Field Experiment," *Management Science* 59(7):1514-1528.

[33] Hwang S., A. Shivdasani, and E. Simintzi, 2019, "Mandating Women on Boards: Evidence from the United States", *working paper.*

[34] Imbens, G. W., and J. D. Angrist, 1994, "Identification and Estimation of Local Average Treatment Effects," *Econometrica* 62(2): 467-475.

[35] Jäger S., B. Schoefer, and J. Heining, 2021, "Labor in the Boardroom," *Quarterly Journal of Economics* 136(2), 669–725.

[36] Jenter, D., T. Schmid, D., Urban, 2019, "Does Board Size Matter?", *working paper.*

[37] Kramer, V.W., A.M. Konrad, S. Erkut, M.J. Hooper, 2007, "Critical mass on corporate boards: why three or more women enhance governance," *Directors Monthly* 31 (2):19-22

[38] Levi, M., K. Li, and F. Zhang, 2014, "Director gender and mergers and acquisitions," *Journal of Corporate Finance* 28, 185–200.

[39] Malmendier, U. and G. Tate, 2005, "CEO Overconfidence and Corporate Investment," *Journal of Finance* 60, 2661–2700.

[40] Matsa, D. A., and A. R. Miller, 2011, "Chipping away at the glass ceiling: Gender spillovers in corporate leadership," *American Economic Review* 101, 635–639.

[41] Matsa, D. A., and A. R. Miller, 2013, "A female style in corporate leadership? Evidence from quotas," *American Economic Journal: Applied Economics* 5, 136–169

[42] von Meyerinck, F., A. Niessen-Ruenzi, M. Schmid, S. D. Solomon, 2020, "As California Goes, So Goes the Nation? Board Gender Quotas and the Legislation of Non-Economic Values", *working paper.*

[43] Nguyen, B. D. and K. M. Nielsen, 2010, "The value of independent directors: Evidence from sudden deaths," *Journal of Financial Economics* 98, 550–567.

[44] Nygaard, K., 2011, "Forced board changes: Evidence from Norway," *working paper.*

[45] Paravisini D., V. Rappoport, P. Schnabl, and D. Wolfenzon, 2014, "Dissecting the Effect of Credit Supply on Trade: Evidence from Matched Credit-Export Data", *Review of Economic Studies* 82(1), 333-359.

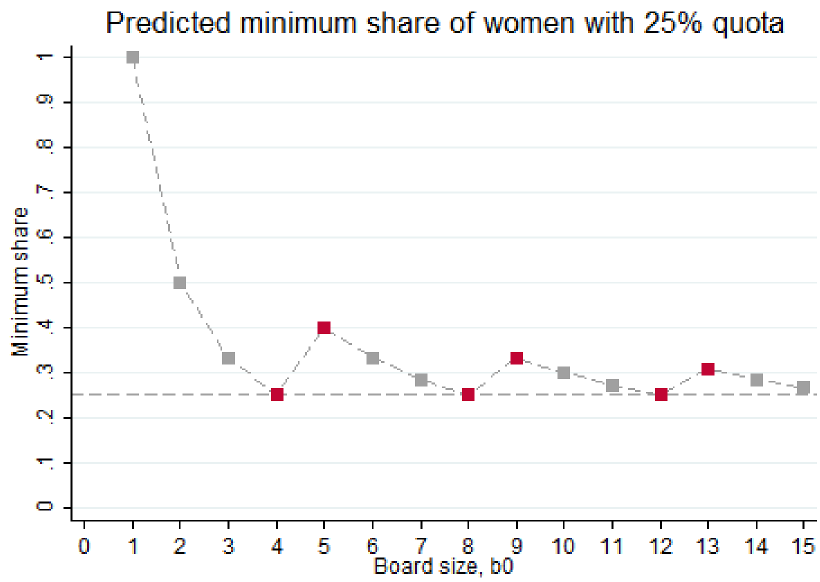
[46] Schwartz-Ziv, M., 2017, "Gender and Board Activeness: The Role of a Critical Mass," *Journal of Financial and Quantitative Analysis* 52(2), pp. 751–780

[47] Seierstad, C., P. Gabaldon, and H. Mensi-Klarbach, 2017, *Gender Diversity in the Boardroom.* Palgrave Macmillan.

[48] Yermack, D., 1996, "Higher Market Valuation of Companies with a Small Board of Directors," *Journal of Financial Economics* 40, 185–211.

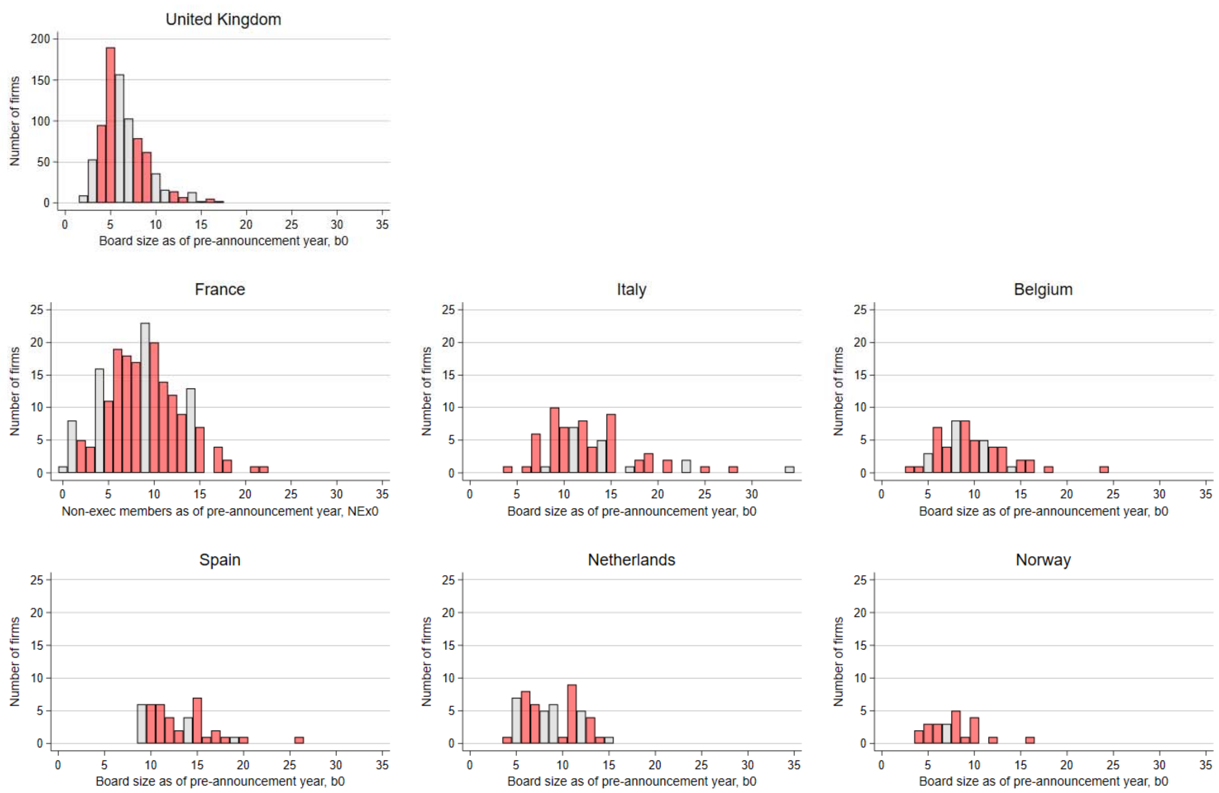
FIGURES AND TABLES

Figure 1. Predicted minimum share, MinShare, as a function of board size



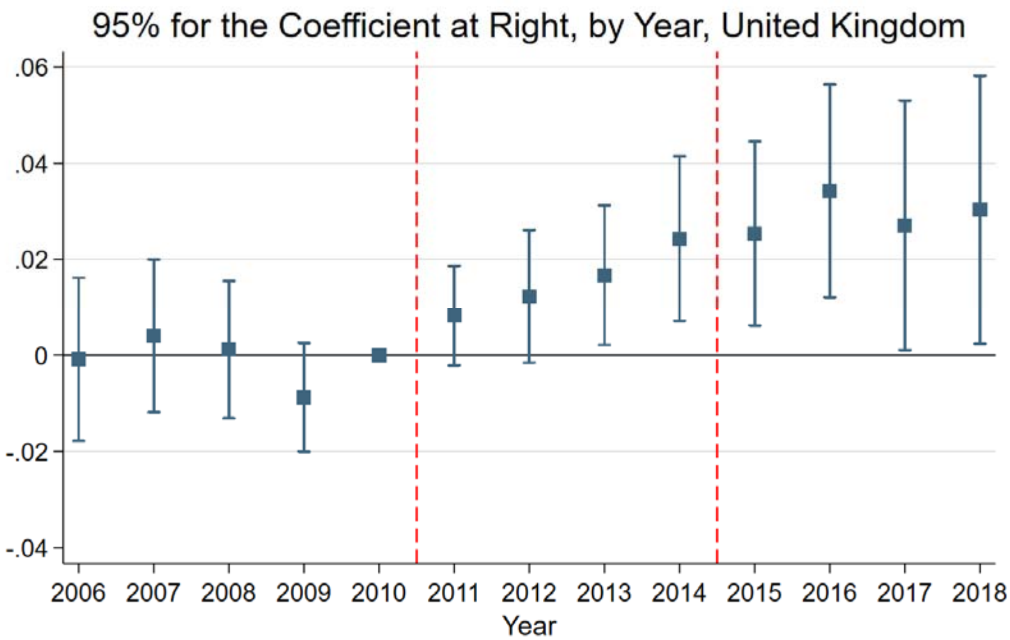
Note: This figure plots the minimum share of women, MinShare, that firms must have to comply with the 25% quota, as a function of board size. The discontinuity samples are highlighted in red.

Figure 2. Distribution of Board Size as of pre-announcement year



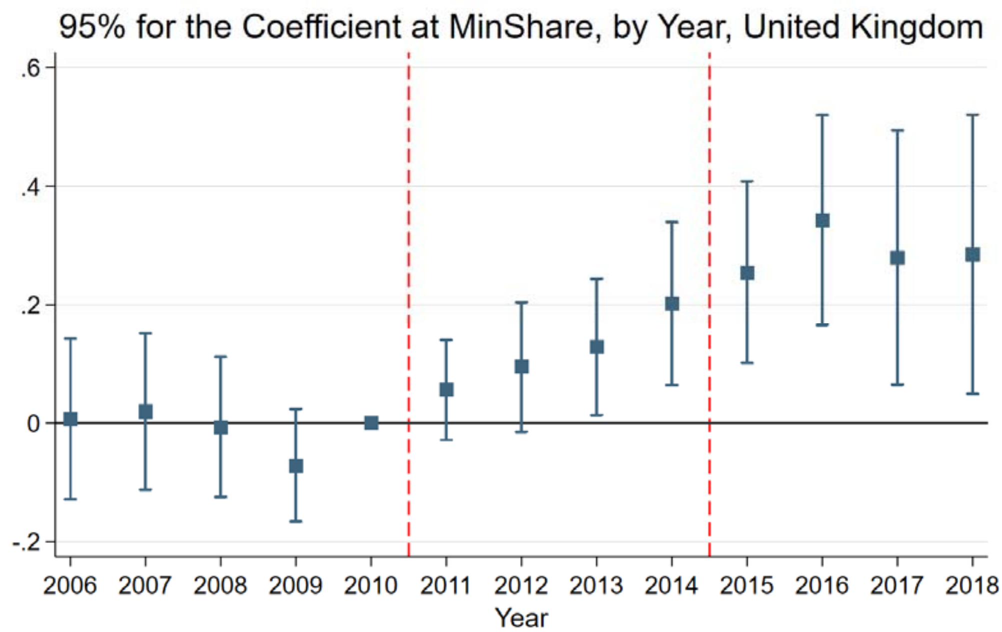
Note: This figure plots the distribution of the pre-quota board size, as of pre-announcement year, for each country separately. The red bars highlight the board sizes of the discontinuity samples analyzed in the paper.

Figure 3. First-stage dynamics – DiD coefficients



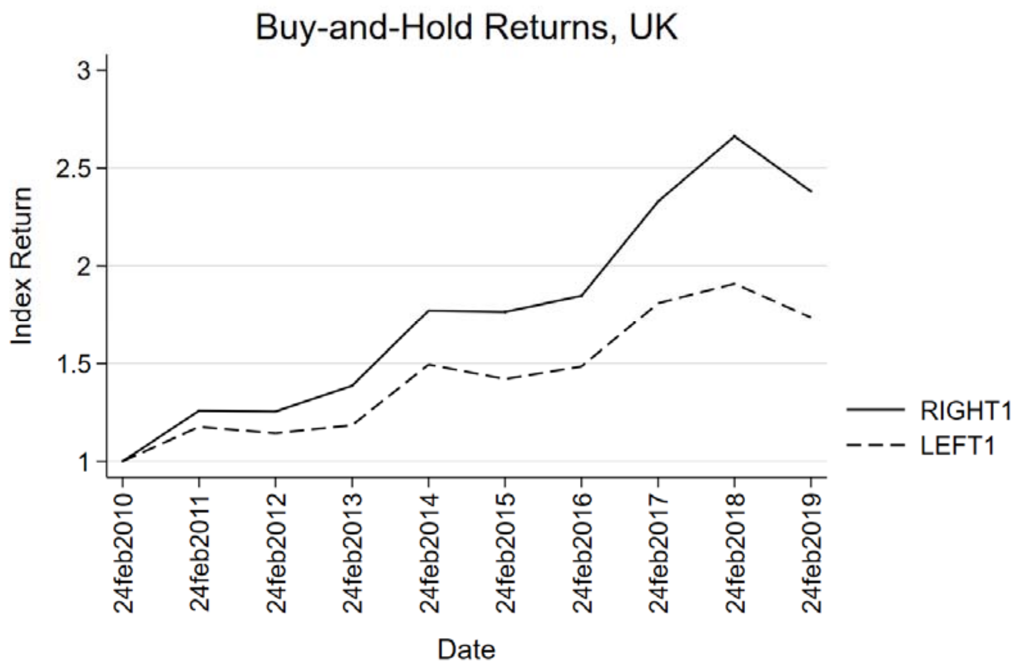
Note: This figure plots the first-stage coefficients from specification (3), for instrument "Right", for the UK. The red dashed lines highlight the years of announcement and compliance, respectively.

Figure 4. First-stage dynamics – DiD coefficients



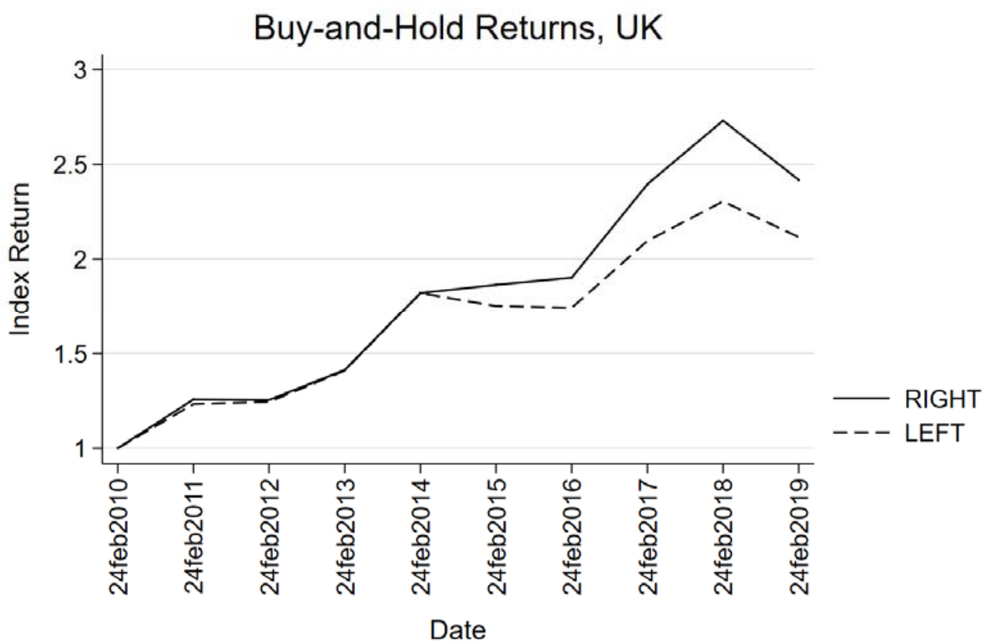
Note: This figure plots the first-stage coefficients from specification (4), for instrument "MinShare", for the UK. The red dashed lines highlight the years of announcement and compliance, respectively.

Figure 5. Buy-and-hold returns in the UK, firms at the largest kink



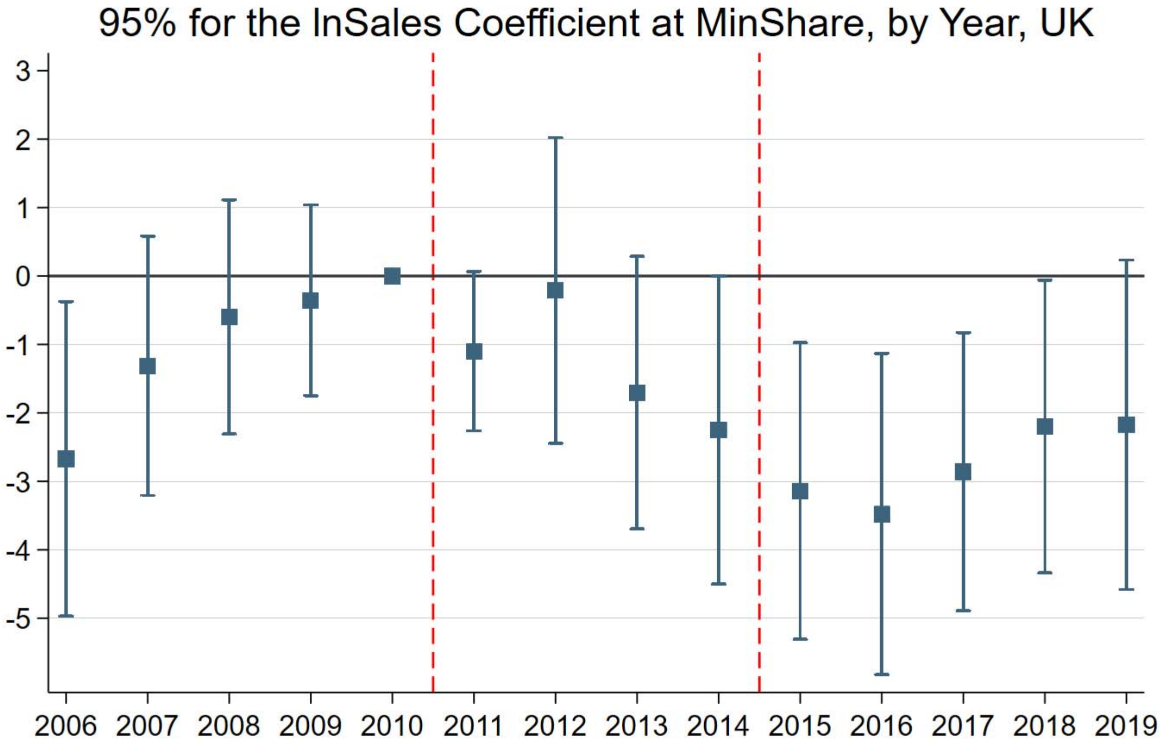
Note: This figure plots buy-and-hold returns that investors would earn by each date if they bought a portfolio of firms to the right (solid line) or to the left (dashed line) of the kink a year before the announcement of the regulation (February 24th, 2010), for the largest kink in the UK.

Figure 6. Buy-and-hold returns in the UK, firms at all kinks



Note: This figure plots buy-and-hold returns that investors would earn by each date if they bought a portfolio of firms to the right (solid line) or to the left (dashed line) of the kink a year before the announcement of the regulation (February 24th, 2010), for all kinks in the UK.

Figure 7. Second-stage dynamics for the Logarithm of Sales – DiD coefficients



Note: This figure plots the dynamics of the second-stage coefficients for the logarithm of sales, based on a yearly regression similar to specification (3), for the UK. The red dashed lines highlight the years of announcement and compliance, respectively.

Table 1. Summary of Quotas and Soft Regulation in European Countries in Sample

Country	Quota or soft regulation in place	Minimum % required	Pre-announcement placebo years in sample	Regulation announcement year	Mandated compliance year	Post-compliance years in sample	N
UK	Self-regulation – from 2012 on the basis of principles of UK CG Code (following the 2011 Lord Davies' recommendation). The recommended target for listed companies in FTSE 100: 25%, by 2015 is applicable to all board members. FTSE 350 companies were recommended setting their own aspirational targets to be achieved.	25%	2008-2010	2011	2015	2015-2018	445
France	Quota of 40% applicable to non-executive directors in large listed and non-listed companies.	40%*	2007-2009	2010	2017	2017-2019	135
Italy	Quota of one-third of each gender for listed companies and state-owned companies to be achieved by 2015.	33%	2008-2010	2011	2015	2015-2018	56
Belgium	Quota for executives and non-executives in state-owned and listed companies - by 2017, in listed SMEs - by 2019.	33%	2008-2010	2011	2017	2017-2019	41
Spain	A gender equality law obliging public companies and IBEX 35-quoted firms with more than 250 employees to attain a minimum 40% share of each gender by 2015.	40%	2004-2006	2007	2015	2015-2018	31
Netherlands	Target of 30% in the boards of large companies by 2016 - "comply or explain" mechanism.	30%	2010-2012	2013	2016	2016-2019	29
Norway	Quota: in February 2002, the government gave a deadline of July 2005 for private listed companies to raise the proportion of women on their boards to 40%. In January 2006 legislation was introduced giving companies a final deadline of January 2008.	40%**	1999-2001	2002	2005	2005-2008	20
						Total:	757

Sources: Davies (2012), European Commission (2016), Seierstad et al (2017)

N is the number of companies in the discontinuity sample as of pre-announcement year that have at least one observation post-compliance.

* As this quota is applicable to non-executive directors only, we consider discontinuity samples that are based on the ex ante number of non-executive directors, rather than the total board size.

** For smaller boards the quota is stated in terms of the number of women, which we account in the analysis. Specifically, boards of less than 3 people should have at least 1 woman, boards of 4-5 people - at least 2 women, boards of 6-8 people - at least 3 women, and larger boards - at least 40% of women.

Table 2. Descriptive Statistics

Note: The table reports the number of observations for all years.

Variable	Mean	STD	N
Financials:			
Total Assets	21 bln	111 bln	5,056
ln (Total Assets)	20.011	2.911	5,056
Market Capitalization	3.97 bln	1.16 bln	5,058
ln (Market Capitalization)	19.572	2.619	5,058
Board Structure and Instrumental Variables:			
Board Size	8.228	3.893	5,059
Board Size as of pre-announcement year (b_{i0})	8.275	4.059	5,059
Share of female directors	0.145	0.155	5,059
Share of female directors as of pre-announcement year	0.063	0.086	5,059
Complier	0.225	0.418	5,059
Predicted minimum required share of women ($MinShare_i$)	0.359	0.078	5,059
Dummy for being to the right of the kink ($Right_i$)	0.525	0.499	5,059
Dependent Variables: Value, Performance, and Other Firm Characteristics			
Tobin's Q	1.848	2.911	5,055
Market Value of Equity to Total Assets	1.294	2.609	5,055
Book Value of Equity to Total Assets	0.477	0.377	5,056
Total Debt to Total Assets	0.192	0.199	5,049
Dividend Payer	0.708	0.455	4,325
Dividend Yield	0.025	0.027	4,298
ln (Sales)	19.262	2.902	4,517
Gross Profit Margin	0.508	0.294	4,024
Operating Profit Margin	-1.054	10.799	4,424
OROA	-0.013	0.348	4,896
ln (Employment)	7.214	2.692	4,215
ln (Labor Productivity)	12.378	1.178	3,879
ln (Average Wage)	10.316	1.581	1,923
Acquisition of a Business	0.798	0.401	2,186
Sale of a Business	0.673	0.469	871
Purchase of Fixed Assets	0.985	0.121	4,093
Sale of Fixed Assets	0.868	0.339	1,940
Beta (Rmrf) - UK only	0.822	0.484	1,420
Beta (SMB) - UK only	0.482	0.533	1,420
Beta (HML) - UK only	0.083	0.603	1,420
Beta (Momentum) - UK only	-0.007	0.325	1,420
Buy-and-Hold Returns - UK only	1.869	2.165	2,629
Dependent Variables: Board Characteristics			
Average age	57.438	4.640	5,057
Average number of qualifications	1.649	0.597	5,059
Average director network size	895.658	717.722	5,059
Average time in company	7.564	4.212	5,059
Share of independent directors	0.484	0.260	5,059
Board meeting attendance	93.846	6.496	1,639
Board specific skills	47.586	23.231	1,848

Table 3. Share of Women, Right and Predicted Minimum Share: First-Stage Results

This table reports the results of estimating the following specification using the OLS framework:

$$Share_{it}=\gamma Post_{ct}Right_i+\lambda_{kct}+\lambda_{sct}+\lambda_i+v_{it} \text{ (columns 1, 2, 4, 5) or}$$

$$Share_{it}=\gamma Post_{ct}MinShare_i+\lambda_{kct}+\lambda_{sct}+\lambda_i+v_{it} \text{ (columns 3 and 6),}$$

where $Share_{it}$ is the fraction of women directors of firm *i* in year *t*, $Post_{ct}$ is the (country-specific) dummy variable that takes value of 1 from announcement year to up to three years afterwards (columns 1 to 3), or from compliance year to three years afterwards (columns 4 to 6), and zero - for the year before announcement and up to three years before that, $Right_i$ is the dummy for being to the right of the kink and $MinShare_i$ is the predicted minimum share of firm i (the instruments, defined in Section 2), measured in the base year, λ_{kct} are kink-year fixed effects (specific to the country), λ_{sct} are industry-year fixed effects (specific to the country), λ_i are firm fixed effects. The base year is the year before announcement. Standard errors are clustered at the firm level and are reported below the coefficients. All columns restrict sample to firms in the discontinuity sample, excluding the potentially non-affected firms (firms with $Share_{i0}$ above the quota). The number of firms and observations excludes singletons.
* indicates 10% significance; ** 5% significance; *** 1% significance.

Panel A: UK	**Post-announcement vs pre-announcement years**			**Post-compliance vs pre-announcement years**		
	1	2	3	4	5	6
$Post_{ct}$ * $Right_i$	0.0216**	0.0161**		0.0647***	0.0303***	
	(0.0102)	(0.00734)		(0.0153)	(0.0112)	
$Post_{ct}$ * $MinShare_i$			0.133**			0.305***
			(0.0600)			(0.0901)
Number of firms	280	444	444	281	445	445
Observations	2,182	3,466	3,466	1,931	3,064	3,064
1st stage F-statistic	4.53	4.80	4.95	17.87	7.38	11.42
Panel B: All countries	**Post-announcement vs pre-announcement years**			**Post-compliance vs pre-announcement years**		
	1	2	3	4	5	6
$Post_{ct}$ * $Right_i$	0.0201**	0.0100*		0.0658***	0.0308***	
	(0.00961)	(0.00599)		(0.0144)	(0.00896)	
$Post_{ct}$ * $MinShare_i$			0.107*			0.319***
			(0.0561)			(0.0840)
Number of firms	360	756	756	361	757	757
Observations	2,769	5,838	5,838	2,418	5,004	5,004
1st stage F-statistic	4.39	2.82	3.62	20.78	11.79	14.38
Kink * Country * Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry * Country * Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample: Largest kink only	Yes			Yes		
Sample: All kinks		Yes	Yes		Yes	Yes

Table 4. Tobin's Q and the Share of Women: Reduced-form and Second-stage Results

This table reports the results of estimating the following specification using the OLS framework:

$$Y_{it}=\gamma Post_{ct}Right_i+\lambda_{kct}+\lambda_{sct}+\lambda_i+v_{it} \text{ (columns 1, 2, 4, 5) or}$$

$$Y_{it}=\gamma Post_{ct}MinShare_i+\lambda_{kct}+\lambda_{sct}+\lambda_i+v_{it} \text{ (columns 3 and 6),}$$

where Y_{it} is Tobin's Q of firm *i* in year *t*, $Post_{ct}$ is the (country-specific) dummy variable that takes value of 1 from compliance year to three years afterwards, and zero - for the year before announcement and up to three years before that, $Right_i$ is the dummy for being to the right of the kink and $MinShare_i$ is the predicted minimum share of firm *i* (the instruments, defined in Section 2), measured in the base year, λ_{kct} are kink-year fixed effects (specific to the country), λ_{sct} are industry-year fixed effects (specific to the country), λ_i are firm fixed effects. The base year is the year before announcement. Standard errors are clustered at the firm level and are reported below the coefficients. All columns restrict sample to firms in the discontinuity sample, excluding the potentially non-affected firms (firms with $Share_{i0}$ above the quota). The number of firms and observations excludes singletons.
* indicates 10% significance; ** 5% significance; *** 1% significance.

Tobin's Q	All countries: Post-compliance vs pre-announcement years			UK: Post-compliance vs pre-announcement years		
	1	2	3	4	5	6
$Post_{ct}$ * $Right_i$	1.011**	0.643**		1.027*	0.799**	
	(0.501)	(0.256)		(0.537)	(0.342)	
$Post_{ct}$ * $MinShare_i$			6.615**			6.947**
			(2.792)			(3.067)
IV-2SLS coefficient	15.36*	20.92**	20.76**	15.88*	26.38*	22.81**
IV-2SLS standard error	(7.982)	(9.710)	(9.817)	(8.726)	(13.92)	(11.49)
Robust Weak-IV AR 95% CI	[0.45, 35.5]	[4.71, 54.04]	[3.69, 50.35]	[-0.41, 38.75]	[4.46, 97.22]	[3.19, 61.74]
Robust Weak-IV AR P-value	0.0438	0.0120	0.0178	0.0558	0.0193	0.0235
1st stage F-statistic	20.78	11.79	14.38	17.87	7.38	11.42
Number of firms	361	757	757	281	445	445
Observations	2,418	5,004	5,004	1,931	3,064	3,064
Kink * Country * Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry * Country * Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample: Largest kink only	Yes			Yes		
Sample: All kinks		Yes	Yes		Yes	Yes

Table 5. Tobin's Q and the Share of Women: Country-Level Analysis

This table reports the results of estimating the following specification using the OLS framework:

$$Y_{it}=\gamma Post_{ct}Right_i+\lambda_{kct}+\lambda_{sct}+\lambda_i+v_{it} \text{ (Panel A) or}$$

$$Y_{it}=\gamma Post_{ct}MinShare_i+\lambda_{kct}+\lambda_{sct}+\lambda_i+v_{it} \text{ (Panel B),}$$

where Y_{it} is Tobin's Q (logarithm of Tobin's Q) of firm *i* in year *t* (Panel A and B, respectively), $Post_{ct}$ is the (country-specific) dummy variable that takes value of 1 from compliance year to three years afterwards, and zero - for the year before announcement and up to three years before that , $Right_i$ is the dummy for being to the right of the kink and $MinShare_i$ is the predicted minimum share of firm *i* (the instruments, defined in Section 2), measured in the base year, λ_{kct} are kink-year fixed effects (specific to the country), λ_{sct} are industry-year fixed effects (specific to the country), λ_i are firm fixed effects. The base year is the year before announcement. Standard errors are clustered at the firm level and are reported below the coefficients. All columns restrict sample to firms in the discontinuity sample, excluding the potentially non-affected firms (firms with $Share_{i0}$ above the quota). The number of firms and observations excludes singletons.
* indicates 10% significance; ** 5% significance; *** 1% significance.

Panel A: $Post_{ct}$ * $Right_i$	**All**	**UK**	**Non-UK**	**FR**	**NL**	**NO**	**BE+IT+SP**
	1	2	3	4	5	6	7
First Stage	0.0337***	0.0296**	0.0640**	0.0910**	0.0649***	0.0423***	0.00217
	(0.0108)	(0.0118)	(0.0249)	(0.0379)	(0.00650)	(0.0123)	(0.0224)
Reduced Form	0.765**	0.816**	0.391**	0.531**	0.169	0.984**	0.0464
	(0.324)	(0.367)	(0.168)	(0.259)	(0.105)	(0.388)	(0.258)
IV-2SLS coefficient	22.68**	27.57*	6.105*	5.836	2.606	23.27***	21.43
IV-2SLS standard error	(11.31)	(15.49)	(3.520)	(3.583)	(1.785)	(2.391)	(322.9)
Robust Weak-IV AR 95% CI	[4.01, 66.21]	[3.45, 125.81]	[0.93, 28.07]	[0.29, 32.45]	[-0.49, 6.71]	[12.34, 26.26]	(-inf, +inf)
Robust Weak-IV AR P-value	0.018	0.026	0.020	0.040	0.106	0.011	0.9370
1st stage F-statistic	9.72	6.29	6.60	5.76	99.74	11.77	0.01
Number of firms	587	419	168	65	23	12	68
Observations	3,923	2,882	1,041	377	157	71	436
Panel B: $Post_{ct}$ * $MinShare_i$	**All**	**UK**	**Non-UK**	**FR**	**NL**	**NO**	**BE+IT+SP**
	1	2	3	4	5	6	7
First Stage	0.328***	0.301***	0.793***	1.133***	0.800***	0.423***	-0.0466
	(0.0881)	(0.0915)	(0.303)	(0.437)	(0.117)	(0.123)	(0.286)
Reduced Form	6.780**	6.866**	5.261**	6.465**	1.770	9.836**	2.168
	(2.956)	(3.120)	(2.084)	(3.151)	(1.306)	(3.878)	(2.992)
IV-2SLS coefficient	20.70**	22.79*	6.630*	5.706*	2.213	23.27***	-46.52
IV-2SLS standard error	(10.07)	(11.80)	(3.661)	(3.458)	(1.572)	(2.391)	(235.0)
Robust Weak-IV AR 95% CI	[3.12, 51.37]	[2.60, 63.79]	[1.38, 28.87]	[0.27, 25.43]	[-1.09, 5.35]	[12.34, 26.26]	(-inf, +inf)
Robust Weak-IV AR P-value	0.021	0.028	0.012	0.040	0.175	0.011	0.8270
1st stage F-statistic	13.83	10.85	6.86	6.72	47.16	11.77	0.03
Number of firms	587	419	168	65	23	12	68
Observations	3,923	2,882	1,041	377	157	71	436
Kink * Country * Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry * Country * Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample: Two largest kinks	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 6. Long-run Buy-and-Hold Returns and the Share of Women: Reduced-form Results for the UK firms

This table reports the results of estimating the following specification using the OLS framework on the sample of UK firms:

$$BH_{it}=\gamma_{2011}D_{2011}Right_i+\gamma_2D_2Right_i+\ldots+\gamma_{2019}D_{2019}Right_i+\lambda_{kt}+\lambda_{st}+\lambda_i+v_{it} ,$$

where BH_{it} is the is the buy-and-hold return of firm *i* in year *t* (i.e. the total return that an investor would earn if she held this stock till year *t*), $Right_i$ is the dummy for being to the right of the kink (as defined before), D_j are the indicator variables for each particular year *j* after announcement, λ_{kt} are kink-year fixed effects, λ_i and λ_{st} are firm and industry-year fixed effects. All returns are measured relative to the year before the announcement (February 24th, 2010), when all D_j are zero and buy-and-hold returns of all firms are mechanically set to 1. The coefficients below report the difference in buy-and-hold returns in each year j, γ_j. Panel A reports the results for the largest kink only, while Panel B considers all kinks. Standard errors are clustered at the industry level and are reported below the coefficients. All columns restrict sample to firms in the discontinuity sample, excluding the potentially non-affected firms (firms with $Share_{i0}$ above the quota). The number of firms and observations excludes singletons. * indicates 10% significance; ** 5% significance; *** 1% significance.

	Long-run Buy-and-Hold Returns, up to year j after announcement, UK firms								
j =	2011	2012	2013	2014	2015	2016	2017	2018	2019
	Panel A: Largest kink only								
$Right_i$	0.124	0.141	0.190*	0.289*	0.352*	0.403*	0.549*	0.703**	0.630**
	(0.0790)	(0.0976)	(0.112)	(0.156)	(0.185)	(0.222)	(0.311)	(0.350)	(0.316)
Number of firms	280	280	280	280	280	280	280	280	280
	Panel B: All kinks								
$Right_i$	0.0782	0.0683	0.0743	0.117	0.230	0.245	0.371*	0.462*	0.346
	(0.0546)	(0.0669)	(0.0829)	(0.122)	(0.143)	(0.159)	(0.225)	(0.268)	(0.238)
Number of firms	443	443	443	443	443	443	443	443	443
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry * Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 7. Long-run Annual Returns and the Share of Women: Reduced-form Results for the UK firms

This table reports the results of estimating the following specification using the OLS framework on the sample of UK firms:

$$R_{it}=\gamma_{2011}Year_{2011}Right_i+\gamma_2 Year_2 Right_i+\ldots+\gamma_{2019}Year_{2019}Right_i+\lambda_{kt}+\lambda_{st}+\lambda_i+v_{it}$$

where R_{it} is the is the total annual return of firm *i* in year *t*, $Right_i$ is the dummy for being to the right of the kink (as defined before), $Year_j$ are the indicator variables for each particular year *j* after announcement, λ_{kt} are kink-year fixed effects, λ_i and λ_{st} are firm and industry-year fixed effects (also absorbing $Right_i$ and all $Year_j$ indicator variables, respectively). The specification is estimated from three years before announcement to the end of the sample. The coefficients below report the difference in annual returns in each year j, γ_j , relative to the pre-announcement average. Panel A reports the results for the largest kink only, while Panel B considers all kinks. Standard errors are clustered at the industry level and are reported below the coefficients. All columns restrict sample to firms in the discontinuity sample, excluding the potentially non-affected firms (firms with $Share_{i0}$ above the quota). The number of firms and observations excludes singletons. * indicates 10% significance; ** 5% significance; *** 1% significance.

	Annual Returns in year j after announcement, UK firms								
j =	2011	2012	2013	2014	2015	2016	2017	2018	2019
	Panel A: Largest kink only								
$Right_i$	0.132**	-0.0120	0.0237	-0.0454	0.0620	0.0001	0.00564	0.0193	-0.0374
	(0.0642)	(0.0508)	(0.0538)	(0.0630)	(0.0518)	(0.0521)	(0.0660)	(0.0562)	(0.0410)
Number of firms	280	280	280	280	280	280	280	280	280
	Panel B: All kinks								
$Right_i$	0.0986**	-0.00440	0.00910	-0.0122	0.0768**	0.0004	0.0354	0.0298	-0.0255
	(0.0446)	(0.0359)	(0.0385)	(0.0442)	(0.143)	(0.0368)	(0.0441)	(0.0382)	(0.0322)
Number of firms	443	443	443	443	443	443	443	443	443
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry * Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 8. Risk Taking and the Share of Women: Reduced-form Results for the UK firms

This table reports the results of estimating the following specification using the OLS framework:

$$Beta_{it}=\gamma Post_t Right_i+\lambda_{kt}+\lambda_{st}+\lambda_i+v_{it}\,,$$

where $Beta_{it}$ is a beta with respect to the 4 factors (with a minimum of 100 days of non-missing observations within that year) of firm *i* in year *t*, $Post_t$ is the dummy variable that takes value of 1 from compliance year to three years afterwards, and zero - for the year before announcement and up to three years before that, $Right_i$ is the dummy for being to the right of the kink (the instrument, defined in Section 2), measured in the base year, λ_{kt} are kink-year fixed effects, λ_{st} are industry-year fixed effects, λ_i are firm fixed effects. The base year is the year before announcement. Standard errors are clustered at the firm level and are reported below the coefficients. All columns restrict sample to firms in the discontinuity sample, excluding the potentially non-affected firms (firms with $Share_{i0}$ above the quota). The number of firms and observations excludes singletons. * indicates 10% significance; ** 5% significance; *** 1% significance.

	Betas, with respect to the 4 factors, UK firms			
	Beta (Rmrf)	Beta (SMB)	Beta (HML)	Beta (Momentum)
	Panel A: Largest kink only			
$Post_t$ * $Right_i$	0.0984	0.100	0.202	0.0771
	(0.115)	(0.101)	(0.135)	(0.105)
Number of firms	157	157	157	157
Observations	618	618	618	618
	Panel B: All kinks			
$Post_t$ * $Right_i$	-0.00719	-0.00945	-0.0715	0.0555
	(0.0547)	(0.0580)	(0.0899)	(0.0405)
Number of firms	302	302	302	302
Observations	1,382	1,382	1,382	1,382
Kink * Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Industry * Year FE	Yes	Yes	Yes	Yes

Table 9. Board Characteristics and the Share of Women: Reduced-form Results

This table reports the results of estimating the following specification using the OLS framework:

$$Y_{it}=\gamma Post_{ct}MinShare_i+\lambda_{kct}+\lambda_{sct}+\lambda_i+v_{it} ,$$

where Y_{it} is the dependent variable of firm *i* in year *t*, $Post_{ct}$ is the (country-specific) dummy variable that takes value of 1 from compliance year to up to three years afterwards, and zero - for the year before announcement and up to three years before that, $MinShare_i$ is the predicted minimum share of firm i (the instruments, defined in Section 2), λ_{kct} are kink-year fixed effects (specific to the country), λ_{sct} are industry-year fixed effects (specific to the country), λ_i are firm fixed effects. The base year is the year before announcement. Standard errors are clustered at the firm level and are reported below the coefficients. All columns restrict sample to firms in the discontinuity sample, excluding the potentially non-affected firms (firms with $Share_{i0}$ above the quota). The number of firms and observations excludes singletons. * indicates 10% significance; ** 5% significance; *** 1% significance.

	Average age	Average number of qualifications	Average network size	Average time in company	Share of independent directors	Average board-specific skills	Average board meetings attendance
	Panel A: All countries						
	1	2	3	4	5	6	7
$Post_{ct}$ * $MinShare_i$	-2.601	-0.116	-410.4	-2.344	-0.0793	35.13	18.12**
	(3.056)	(0.375)	(465.0)	(2.548)	(0.127)	(39.70)	(8.389)
Number of firms	766	766	766	766	766	338	308
Observations	5,055	5,055	5,055	5,055	5,055	1,815	1,605
	Panel B: UK						
	1	2	3	4	5	6	7
$Post_{ct}$ * $MinShare_i$	-1.236	-0.0738	-453.6	-1.423	-0.0389	31.01	15.54*
	(3.252)	(0.410)	(513.7)	(2.735)	(0.137)	(45.00)	(8.749)
Number of firms	445	445	445	445	445	142	143
Observations	3,064	3,064	3,064	3,064	3,064	851	854
Kink * Country * Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry * Country * Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 10. Q Decomposition, Leverage, and Operating Performance: Reduced-form Results

This table reports the results of estimating the following specification using the OLS framework:

$$Y_{it}=\gamma Post_{ct}MinShare_i+\lambda_{kct}+\lambda_{sct}+\lambda_i+v_{it} ,$$

where Y_{it} is the dependent variable of firm *i* in year *t*, $Post_{ct}$ is the (country-specific) dummy variable that takes value of 1 from compliance year to three years afterwards, and zero - for the year before announcement and up to three years before that, $MinShare_i$ is the predicted minimum share of firm *i* (the instrument, defined in Section 2), measured in the base year, λ_{kct} are kink-year fixed effects (specific to the country), λ_{sct} are industry-year fixed effects (specific to the country), λ_i are firm fixed effects. The base year is the year before announcement. Standard errors are clustered at the firm level and are reported below the coefficients. All columns restrict sample to firms in the discontinuity sample, excluding the potentially non-affected firms (firms with $Share_{i0}$ above the quota). The number of firms and observations excludes singletons. * indicates 10% significance; ** 5% significance; *** 1% significance.

	Market Value of Equity / Total Assets	Book Value of Equity / Total Assets	Total Debt / Total Assets	Dividend Yield	ln Total Assets	OROA
	Panel A: All countries					
	1	2	3	4	5	6
$Post_{ct}$ * $MinShare_i$	5.349**	-0.265	-0.0464	-0.0334**	-2.275***	-0.926***
	(2.700)	(0.278)	(0.135)	(0.0161)	(0.722)	(0.314)
Number of firms	766	766	766	754	766	742
Observations	5,055	5,055	5,048	4,292	5,055	4,895
	Panel B: UK					
	1	2	3	4	5	6
$Post_{ct}$ * $MinShare_i$	5.362*	-0.283	-0.0291	-0.0295*	-2.395***	-1.055***
	(2.808)	(0.309)	(0.148)	(0.0173)	(0.793)	(0.348)
Number of firms	445	445	445	442	445	440
Observations	3,064	3,064	3,059	2,690	3,064	3,029
Kink * Country * Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry * Country * Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 11. Investments and the Share of Women: Reduced-form Results

This table reports the results of estimating the following specification using the OLS framework:

$$Y_{it}=\gamma Post_{ct}MinShare_i+\lambda_{kct}+\lambda_{sct}+\lambda_i+\nu_{it} ,$$

where Y_{it} is the dependent variable of firm *i* in year *t*, $Post_{ct}$ is the (country-specific) dummy variable that takes value of 1 from compliance year to three years afterwards, and zero - for the year before announcement and up to three years before that, $MinShare_i$ is the predicted minimum share of firm *i* (the instrument, defined in Section 2), measured in the base year, λ_{kct} are kink-year fixed effects (specific to the country), λ_{sct} are industry-year fixed effects (specific to the country), λ_i are firm fixed effects. The base year is the year before announcement. Standard errors are clustered at the firm level and are reported below the coefficients. All columns restrict sample to firms in the discontinuity sample, excluding the potentially non-affected firms (firms with $Share_{i0}$ above the quota). The number of firms and observations excludes singletons. * indicates 10% significance; ** 5% significance; *** 1% significance.

	Acquisition of Business	Sale of Business	Purchase of Fixed Assets	Sale of Fixed Assets
	Panel A: All countries			
	1	2	3	4
$Post_{ct}$ * $MinShare_i$	-1.039**	0.364	-0.182*	-0.0950
	(0.476)	(0.799)	(0.107)	(0.318)
Number of firms	463	207	661	372
Observations	2,137	805	4,089	1,880
	Panel B: UK			
	1	2	3	4
$Post_{ct}$ * $MinShare_i$	-1.303**	0.419	-0.178	-0.0803
	(0.521)	(0.805)	(0.122)	(0.323)
Number of firms	251	152	353	267
Observations	1,167	593	2,325	1,428
Kink * Country * Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Industry * Country * Year FE	Yes	Yes	Yes	Yes

Table 12. Operating Performance and the Share of Women: Reduced-form Results

This table reports the results of estimating the following specification using the OLS framework:

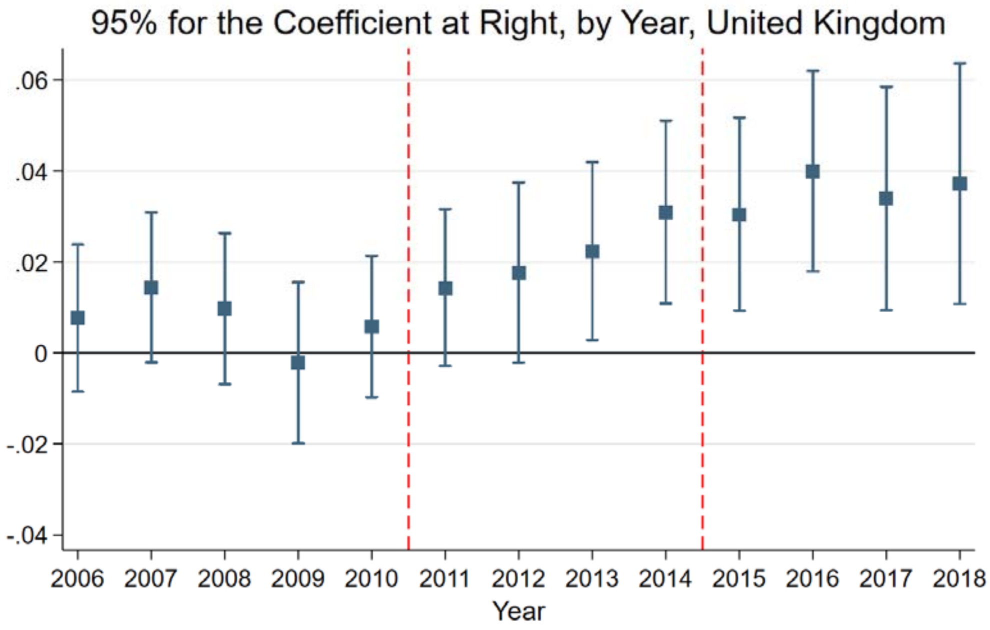
$$Y_{it}=\gamma Post_{ct}MinShare_i+\lambda_{kct}+\lambda_{sct}+\lambda_i+v_{it} ,$$

where Y_{it} is the dependent variable of firm *i* in year *t*, $Post_{ct}$ is the (country-specific) dummy variable that takes value of 1 from compliance year to three years afterwards, and zero - for the year before announcement and up to three years before that, $MinShare_i$ is the predicted minimum share of firm *i* (the instrument, defined in Section 2), measured in the base year, λ_{kct} are kink-year fixed effects (specific to the country), λ_{sct} are industry-year fixed effects (specific to the country), λ_i are firm fixed effects. The base year is the year before announcement. Standard errors are clustered at the firm level and are reported below the coefficients. All columns restrict sample to firms in the discontinuity sample, excluding the potentially non-affected firms (firms with $Share_{i0}$ above the quota). The number of firms and observations excludes singletons. * indicates 10% significance; ** 5% significance; *** 1% significance.

	ln Sales	Sales / Total Assets	Operating Expenses / Total Assets	Gross Profit Margin	Operating Profit Margin	ln Employment	ln Labor Productivity	ln Average Wage
				Panel A: All countries				
	1	2	3	4	5	6	7	8
$Post_{ct}$ * $MinShare_i$	-3.043***	-0.818**	0.432	-0.256	0.395	-2.039***	-0.420	-0.868
	(0.766)	(0.324)	(0.493)	(0.165)	(7.606)	(0.755)	(0.576)	(1.473)
Number of firms	725	742	742	667	711	691	643	359
Observations	4,511	4,894	4,894	4,011	4,418	4,210	3,868	1,889
				Panel B: UK				
	1	2	3	4	5	6	7	8
$Post_{ct}$ * $MinShare_i$	-3.318***	-0.919***	0.498	-0.219	-0.001	-2.448***	-0.364	-2.317
	(0.856)	(0.356)	(0.545)	(0.187)	(8.724)	(0.885)	(0.690)	(2.307)
Number of firms	423	440	440	389	419	372	343	134
Observations	2,649	3,028	3,028	2,336	2,622	2,287	2,073	681
Kink * Country * Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry * Country * Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

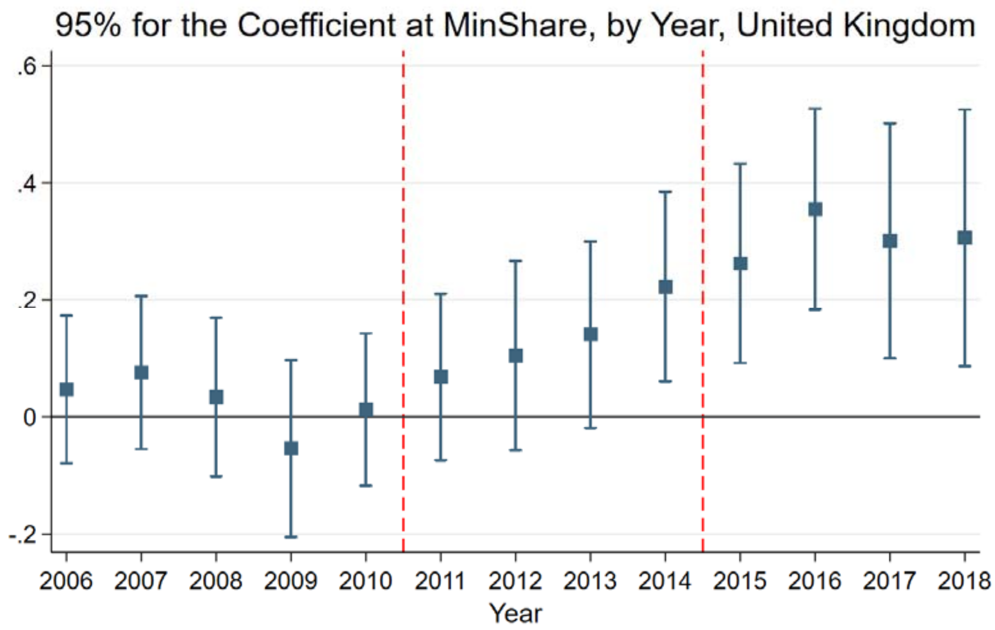
INTERNET APPENDIX

Appendix Figure 1. First-stage dynamics – cross-sectional coefficients



Note: This figure plots the first-stage coefficients from specification (3) without firm fixed effects, for instrument "Right", for the UK. The red dashed lines highlight the years of announcement and compliance, respectively.

Appendix Figure 2. First-stage dynamics – cross-sectional coefficients



Note: This figure plots the first-stage coefficients from specification (4) without firm fixed effects, for instrument "MinShare", for the UK. The red dashed lines highlight the years of announcement and compliance, respectively.

Appendix Table 1. Share of Women and the Instruments: Placebo Past Trends

This table reports the results of estimating the following specification using the OLS framework:

$$\Delta Share_{it}=\gamma Right_i+\lambda_{kc}+\lambda_{sc}+v_{it} \text{ (columns 1, 2, 4, 5) or}$$

$$\Delta Share_{it}=\gamma MinShare_i+\lambda_{kc}+\lambda_{sc}+v_{it} \text{ (columns 3 and 6),}$$

where $\Delta Share_{it}$ is the yearly change in the fraction of women directors of firm *i* in year *t*, $Right_i$ is the dummy for being to the right of the kink, $MinShare_i$ is the predicted minimum share of firm *i* (the instruments, defined in Section 2), all as of the year before announcement, λ_{kc} are kink fixed effects (specific to the country), λ_{sc} are industry fixed effects (specific to the country). Standard errors are clustered at the firm level and are reported below the coefficients. All columns restrict sample to firms in the discontinuity sample, excluding the potentially non-affected firms (firms with $Share_{i0}$ above the quota). All regressions are estimated using all available data before announcement (up to 10 years before). The number of firms and observations excludes singletons.
* indicates 10% significance; ** 5% significance; *** 1% significance.

	Instruments used in the paper					
Share of female directors	**All countries**			**UK**		
	1	2	3	4	5	6
$Right_i$	-0.00225	-0.00109		-0.00133	-0.000433	
	(0.00252)	(0.00127)		(0.00269)	(0.00154)	
$MinShare_i$			-0.00707			-0.00333
			(0.0131)			(0.0140)
Number of firms	364	771	771	289	455	455
Observations	2,088	5,168	5,168	1,659	3,048	3,048
Kink * Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry * Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample: Largest kink only	Yes			Yes		
Sample: All kinks		Yes	Yes		Yes	Yes

Appendix Table 2. Tobin's Q and the Instruments: Placebo Past Trends

This table reports the results of estimating the following specification using the OLS framework:

$$\Delta Y_{it}=\gamma Right_i+\lambda_{kc}+\lambda_{sc}+\nu_{it} \text{ (columns 1, 2, 4, 5) or}$$

$$\Delta Y_{it}=\gamma MinShare_i+\lambda_{kc}+\lambda_{sc}+\nu_{it} \text{ (columns 3 and 6) or}$$

$$\Delta Y_{it}=\gamma Woman_i+\lambda_{sc}+\nu_{it} \text{ (columns 7 and 8),}$$

$$\Delta Y_{it}=\gamma Share_i+\lambda_{sc}+\nu_{it} \text{ (columns 9 and 10),}$$

where ΔY_{it} is the yearly change in Tobin's Q of firm *i* in year *t*, $Right_i$ is the dummy for being to the right of the kink, $MinShare_i$ is the predicted minimum share of firm *i* (the instruments, defined in Section 2), $Woman_i$ is the dummy for having at least one woman, $Share_i$ is the share of women, all as of the year before announcement, λ_{kc} are kink fixed effects (specific to the country), λ_{sc} are industry fixed effects (specific to the country). Standard errors are clustered at the firm level and are reported below the coefficients. Columns 1 to 6 restrict sample to firms in the discontinuity sample, excluding the potentially non-affected firms (firms with $Share_{i0}$ above the quota). Columns 7 to 10 consider all firms. All regressions are estimated using all available data before announcement (up to 10 years before). The number of firms and observations excludes singletons. * indicates 10% significance; ** 5% significance; *** 1% significance.

Tobin's Q	Instruments used in the paper						Not used			
	All countries			UK			All countries		UK	
	1	2	3	4	5	6	7	8	9	10
$Right_i$	-0.0398	0.0717		-0.0267	0.0859					
	(0.276)	(0.135)		(0.298)	(0.177)					
$MinShare_i$			0.473			0.496				
			(1.491)			(1.655)				
$Woman_i$							0.133**		0.153*	
							(0.0670)		(0.0881)	
$Share_i$								0.625**		0.688**
								(0.263)		(0.338)
Number of firms	371	784	784	291	457	457	1311	1311	850	850
Observations	2,314	5,533	5,533	1,832	3,290	3,290	9,441	9,441	6,241	6,241
Kink * Country FE	Yes	Yes	Yes	Yes	Yes	Yes				
Industry * Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample: Largest kink only	Yes			Yes						
Sample: All kinks		Yes	Yes		Yes	Yes				
Sample: All board sizes							Yes	Yes	Yes	Yes

Appendix Table 3. Risk Taking and the Share of Women: Reduced-form Results for the UK firms

This table reports the results of estimating the following specification using the OLS framework:

$$Beta_{it}=\gamma Post_t Right_i+\lambda_{kct}+\lambda_{sct}+\lambda_i+v_{it} ,$$

where $Beta_{it}$ is a beta with respect to the 4 factors (with a minimum of 100 days of non-missing observations within that year) of firm *i* in year *t*, $Post_t$ is the dummy variable that takes value of 1 from announcement year to three years afterwards, and zero - for the year before announcement and up to three years before that, $Right_i$ is the dummy for being to the right of the kink (the instrument, defined in Section 2), measured in the base year, λ_{kt} are kink-year fixed effects, λ_{st} are industry-year fixed effects, λ_i are firm fixed effects. The base year is the year before announcement. Standard errors are clustered at the firm level and are reported below the coefficients. All columns restrict sample to firms in the discontinuity sample, excluding the potentially non-affected firms (firms with $Share_{i0}$ above the quota). The number of firms and observations excludes singletons. * indicates 10% significance; ** 5% significance; *** 1% significance.

	Betas, with respect to the 4 factors, UK firms			
	Beta (Rmrf)	Beta (SMB)	Beta (HML)	Beta (Momentum)
	Panel A: Largest kink only			
$Post_t$ * $Right_i$	0.106	0.108	-0.0844	0.113
	(0.114)	(0.126)	(0.217)	(0.0917)
Number of firms	149	149	149	149
Observations	697	697	697	697
	Panel B: All kinks			
$Post_t$ * $Right_i$	0.0161	0.0306	-0.0713	0.0632*
	(0.0489)	(0.0587)	(0.0848)	(0.0359)
Number of firms	293	293	293	293
Observations	1,581	1,581	1,581	1,581
Kink * Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Industry * Year FE	Yes	Yes	Yes	Yes

Appendix Table 4. Board Size: Reduced-form Results

This table reports the results of estimating the following specification using the OLS framework:

$$Y_{it}=\gamma Post_{ct}Right_i+\lambda_{kct}+\lambda_{sct}+\lambda_i+v_{it} \text{ (columns 1, 2) or}$$

$$Y_{it}=\gamma Post_{ct}MinShare_i+\lambda_{kct}+\lambda_{sct}+\lambda_i+v_{it} \text{ (column 3),}$$

where Y_{it} is Board size of firm *i* in year *t*, $Post_{ct}$ is the (country-specific) dummy variable that takes value of 1 from compliance year to three years afterwards, and zero - for the year before announcement and up to three years before that, $Right_i$ is the dummy for being to the right of the kink and $MinShare_i$ is the predicted minimum share of firm *i* (the instruments, defined in Section 2), measured in the base year, λ_{kct} are kink-year fixed effects (specific to the country), λ_{sct} are industry-year fixed effects (specific to the country), λ_i are firm fixed effects. The base year is the year before announcement. Standard errors are clustered at the firm level and are reported below the coefficients. All columns restrict sample to firms in the discontinuity sample, excluding the potentially non-affected firms (firms with $Share_{i0}$ above the quota). The number of firms and observations excludes singletons. * indicates 10% significance; ** 5% significance; *** 1% significance.

	Board Size		
	1	2	3
	Panel A: All countries		
$Post_{ct}$ * $Right_i$	-0.218	-0.383***	
	(0.153)	(0.138)	
$Post_{ct}$ * $MinShare_i$			-2.677**
			(1.041)
Number of firms	361	766	766
Observations	2,418	5,055	5,055
	Panel B: UK		
$Post_{ct}$ * $Right_i$	-0.214	-0.371***	
	(0.155)	(0.135)	
$Post_{ct}$ * $MinShare_i$			-2.480**
			(1.012)
Number of firms	281	445	445
Observations	1,931	3,064	3,064
Kink * Country * Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Industry * Country * Year FE	Yes	Yes	Yes
Sample: Largest kink only	Yes		
Sample: All kinks		Yes	Yes

Appendix Table 5. Tobin's Q and the Share of Women: Reduced-form and Second-stage Results

This table reports the results of estimating the following specification using the OLS framework:

$$Y_{it}=\gamma Post_{ct}Right_i+BoardSize_{it}+\lambda_{kct}+\lambda_{sct}+\lambda_i+v_{it} \text{ (columns 1, 2, 4, 5) or}$$

$$Y_{it}=\gamma Post_{ct}MinShare_i+BoardSize_{it}+\lambda_{kct}+\lambda_{sct}+\lambda_i+v_{it} \text{ (columns 3 and 6),}$$

where Y_{it} is Tobin's Q of firm *i* in year *t*, $Post_{ct}$ is the (country-specific) dummy variable that takes value of 1 from from compliance year to three years afterwards, and zero -- for the year before announcement and up to three years before that, $Right_i$ is the dummy for being to the right of the kink and $MinShare_i$ is the predicted minimum share of firm *i* (the instruments, defined in Section 2), measured in the base year, $BoardSize_{it}$ is the board size of firm *i* in year *t*, λ_{kct} are kink-year fixed effects (specific to the country), λ_{sct} are industry-year fixed effects (specific to the country), λ_i are firm fixed effects. The base year is the year before announcement. Standard errors are clustered at the firm level and are reported below the coefficients. All columns restrict sample to firms in the discontinuity sample, excluding the potentially non-affected firms (firms with $Share_{i0}$ above the quota). The number of firms and observations excludes singletons. * indicates 10% significance; ** 5% significance; *** 1% significance.

Tobin's Q	**All countries: Post-compliance vs pre-announcement years**			**UK: Post-compliance vs pre-announcement years**		
	1	2	3	4	5	6
$Post_{ct}$ * $Right_i$	1.040**	0.650**		1.055*	0.824**	
	(0.507)	(0.260)		(0.544)	(0.351)	
$Post_{ct}$ * $MinShare_i$			6.581**			7.106**
			(2.787)			(3.125)
$BoardSize_{it}$	0.133	0.0388	0.0381	0.129	0.0655	0.0640
	(0.111)	(0.0360)	(0.0357)	(0.126)	(0.0657)	(0.0652)
IV-2SLS coefficient	15.43*	19.22**	19.60**	15.87*	23.75**	21.28**
IV-2SLS standard error	(7.893)	(8.656)	(9.165)	(8.601)	(11.85)	(10.45)
Robust Weak-IV AR 95% CI	[0.71, 35.15]	[4.28, 45.26]	[3.42, 45.59]	[-0.18, 38.09]	[4.05, 67.99]	[3.03, 52.92]
Robust Weak-IV AR P-value	0.0404	0.0122	0.0182	0.0525	0.0190	0.0230
1st stage F-statistic	21.80	14.77	16.48	18.88	10.07	14.01
Number of firms	361	757	757	281	445	445
Observations	2,418	5,004	5,004	1,931	3,064	3,064
Kink * Country * Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry * Country * Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample: Largest kink only	Yes			Yes		
Sample: All kinks		Yes	Yes		Yes	Yes