Comparative Political Studies

From Open to Secret Ballot: Vote Buying and Modernization

Journal:	*Comparative Political Studies*
Manuscript ID	CPS-15-0259.R1
Manuscript Type:	Original Article
Keywords:	Democratization and Regime Change, Political Economy

Abstract: The secret ballot is one of the cornerstones of democracy. We contend that the historical process of modernization caused the switch from open to secret ballot with the underlying mechanism being that income growth, urbanization, and rising education standards undermined vote markets. We undertake event history studies of ballot reform in Western Europe and the US states during the 19th and 20th centuries to establish that modernization was systematically related to ballot reform. We study electoral turnout before and after ballot reform amongst the US states and British parliamentary constituencies to substantiate the hypothesis that modernization reduced the volume of trade in the vote market.

From Open to Secret Ballot: Vote Buying and Modernization[1]

Toke S. Aidt, University of Cambridge and CESifo, Munich[2]

Peter S. Jensen, University of Southern Denmark[3]

October 18, 2015

The secret ballot is one of the cornerstones of democracy. We contend that the historical process of modernization caused the switch from open to secret ballot with the underlying mechanism being that income growth, urbanization, and rising education standards undermined vote markets. We undertake event history studies of ballot reform in Western Europe and the US states during the 19th and 20th centuries to establish that modernization was systematically related to ballot reform. We study electoral turnout before and after ballot reform amongst the US states and British parliamentary constituencies to substantiate the hypothesis that modernization reduced the volume of trade in the vote market.

Key words: Secret ballot, modernization, electoral turnout, democratization.
JEL classification: D7, P16.

[1] Acknowledgements: We would like to thank Julia Shvets, Miriam Golden, Jo Andrews, James Robinson, Thomas Barnebeck Andersen, Christian Bjørnskov, Carl-Johan Dalgaard, Casper Worm Hansen, Ingrid Henriksen, Lars Lønstrup, Martin Paldam, Paul Sharp, Linda Veiga, Jacob Gerner Hariri, Philipp Schröder, Daniel Ziblatt, Jørgen Elklit, Jan Teorell, Chris Kam, Fabrice Lehoucq, Isabela Mares, John Gerring and participants in the Harvard workshop "The Causes and Consequences of Electoral Fraud: Lessons from the History of Established Democracies" for helpful comments and useful suggestions on preliminary drafts of this paper. We also thank Samuel Berlinski, Torun Dewan and James Fearon, for sharing data with us. This paper was partly written while Toke Aidt was visiting the Institute for Quantitative Social Science at Harvard University and, subsequently, the CESifo in Munich. The hospitality of these institutes is much appreciated. The paper was presented at the 2nd World Congress of the Public Choice Society and European Public Choice Society in Miami in March 2012. The usual qualifier applies.
[2] Corresponding Author: Faculty of Economics, University of Cambridge, Cambridge CB3 9DD, United Kingdom. E-mail: toke.aidt@econ.cam.ac.uk. Phone: +44 1223 335231. Fax: +44 1223 335475.
[3] Department of Business and Economics, Campusvej 55, 5230 Odense M, Denmark. E-mail: psj@sam.sdu.dk.

From Open to Secret Ballot: Vote Buying and Modernization

The secret ballot is one of the cornerstones of democracy. We contend that the historical process of modernization caused the switch from open to secret ballot with the underlying mechanism being that income growth, urbanization, and rising education standards undermined vote markets. We undertake event history studies of ballot reform in Western Europe and the US states during the 19th and 20th centuries to establish that modernization was systematically related to ballot reform. We study electoral turnout before and after ballot reform amongst the US states and British parliamentary constituencies to substantiate the hypothesis that modernization reduced the volume of trade in the vote market.

Key words: Secret ballot, modernization, electoral turnout, democratization.
JEL classification: D7, P16.

1 Introduction

One of the most striking facts in comparative political economy is the positive correlation between national income and democracy. Since the first statistical evidence was unearthed in the 1950s by Seymour M. Lipset in his influential paper "Social Requisites of Democracy: Economic Development and Political Legitimacy" (Lipset, 1959), a lively debate amongst political scientists, sociologists, and economists regarding the correct interpretation of this correlation has raged. Lipset (1959, p. 86) himself interprets, in what has subsequently become known as modernization theory, the correlation as a unidirectional causal relationship from economic development to democracy or as he puts it "economic development involving industrialization, urbanization, higher educational standards and a steady increase in the overall wealth of the society is a basic condition sustaining democracy". This interpretation is controversial.[1]

We propose a new perspective on the modernization debate that we believe will help explore the boundaries of the theory in a more constructive way. We begin by observing that democracy is a package of institutions. This observation is neither new nor novel and most

[1] Amongst those skeptical, we count Moore (1966), Przeworski and Limongi (1997), and Acemoglu et al. (2008), while Barro (1999), Gundlach and Paldam (2009) and Boix (2011) present evidence consistent with Lipset's original interpretation. Others, e.g., Rueschemeyer et al. (1993), Ansell and Samuels (2015), Miller (2013), and Treisman (2015) have emphasized different mechanisms than Lipset (1959).

writers follow Dahl (1971) and make a distinction between different aspects of democracy. Yet, the modernization debate centers on the causal relationship between GDP per capita and *composite* indices of democracy. It is, therefore, either assumed that democratization is an all or nothing choice–a view that is contradicted by the historical record–or that all the sub-components of the overall package are equally likely (or unlikely) to be causally driven by modernization. The alternative we propose is that modernization may be causally linked to specific sub-components of the overall package of democratic institutions without necessarily governing, in a causal sense, the evolution of the overall package or all of its parts.

We focus on the secret ballot for three related reasons.[2] First, the secret ballot is regarded as one of the cornerstones of fair elections (e.g., Rokkan, 1961; Elklit, 2000; Alvarez et al., 1996). Baland and Robinson (2007, p. 140) note that "the introduction of political institutions that stop corruption and vote buying, such as the Australian ballot, appear to be as significant a step in the process of political development as the construction of electoral democracy itself." Against this background, gaining a better understanding of how the secret ballot came about is important in itself.[3] Second, there is a straightforward causal mechanism that links modernization to the secret ballot. The mechanism operates through the vote market. Vote markets thrive under open or semi-open voting because this allows the buyer of a vote to verify that the seller kept his part of the bargain. Modernization tends to erode social control and the scope for economic sanction, to improve outside options for ordinary voters, or to undermine old norms of social deference. Income growth also tends to increase the price of a vote. All of these forces combine to make vote buying less economical and the importance of vote markets is diminished. The defenders of the open ballot then become less stout defenders and ballot reform becomes more likely. Third, as we discuss in more detail below, a large body of research suggests that it is modernization and development that destroy clentelism in modern democracies. Historically, similar processes were in operation (e.g., Stokes et al., 2013) and the demise of the open vote and the introduction of secrecy in voting were a potentially important part of this process.

We begin the analysis by formalizing our reasoning in a novel rational choice model of ballot reform. The model demonstrates why modernization undermines the vote market and how

[2] Boix (1999), Blais et al. (2004), and Andrews and Jackman (2005) study the adoption of proportional representation, Przeworski (2008, 2009), Congleton (2011), Aidt and Franck (2015), Aidt and Jensen (2014) and Aidt and Leon (2015) study the causes of suffrage reform. The consequences (but not the causes) of secret ballot have been studied by Anderson and Tollison (1990), Heckelman (1995), and Aidt and Jensen (2009).

[3] See also Przeworski (2010).

this triggers the secret ballot. The logic is that modernization, while not necessarily eliminating vote buying altogether, reduces the importance of vote markets for those groups who benefit from their existence**.** This happens because the price of a vote is pushed up as voters become richer or become less willing to conform with norms of social deference. As a consequence of this, the hand of those groups which stand to gain from secret ballot is strengthened, while the hand of the groups that benefit from vote buying under open ballot is weakened. This makes ballot reform more likely and when the secret ballot is introduced, the scope for vote buying is further reduced (in the model eliminated entirely). The model, therefore, predicts i) that modernization makes secret ballot more likely and ii) that it gradually undermines (but does not eliminate) vote markets in the lead up to the introduction of secret ballot at which point vote markets are further undermined**.** To support these predictions empirically, we marshal two types of evidence. The first type is based on event history studies of the adoption of the secret ballot. This research design enables us to ask whether modernization–higher income levels, urbanization, and higher education standards–predicts the timing of ballot reform in two different historical samples: Western Europe plus English-speaking off-shoots (1820-1913) and US states (1840-1950). In both cases, we find strong evidence that modernization affected the timing of the secret ballot. In contrast, the same modernization variables cannot predict the timing of reforms that extended the suffrage to broader segments of the male population (Przeworski, 2009; Aidt and Jensen, 2014). The second type of evidence delves deeper into the underlying causal mechanism. With the secret ballot, the vote loses (much of) its value as a tradable commodity and voters have one less reason to vote. The implied drop in turnout can, therefore, be taken as an indicator of the importance of the vote market under open ballot. We can, then, ask whether the fall in turnout is smaller in places where modernization has progressed more, as one would expect if modernization encourages ballot reform by making electoral corruption less economical. We investigate this in a sample of US states (1870-1950) and in a sample of parliamentary constituencies of Great Britain in the election before and after the Ballot Act of 1872. In both contexts, we find evidence consistent with the proposed causal mechanism.

Our analysis speaks to a large literature on clientelism. A common theme of this literature is that structural forces, such as economic growth, industrialization, urbanization and expanding electorates, tend to undermine politics based on the exchange of a citizen's vote for payments or continuing access to employment, goods and services and tend to encourage political parties to

establish other modes of distributive politics or to develop programmatic policy platforms. Kitschelt and Wilkinson (2007) emphasize how such structural forces interact with party competition to generate complex nonlinear dynamics. Stokes et al. (2013) emphasize monitoring and incentive problems in the relationship between party leaders and the intermediaries or brokers they need to employ in order to deliver private benefits to particular voters and to keep track of the holding-voters-to-account side of the vote market. Keefer and Vlaicu (2008) emphasize that vote buying emerges when political parties are unable to commit to programmatic policy platforms or lack the technologies to communicate them to large groups of voters. In all cases, modernization tends to make vote buying uneconomical and to increase the relative return to alternative electoral strategies. As in the broader literature on clientelism, our new theory of ballot reform emphasizes that modernization—income growth and education in particular—and franchise extension undermine vote markets. We argue that this generates a causal link running from modernization to ballot reform. This, in turn, generates particular patterns in electoral turnout that enables a more refined test of the theory. Our empirical finding—modernization was important for the demise of open voting—not only provides systematic, statistical evidence consistent with our particular model but also speaks to the broader literature on modernization and clientelism. The statistical analysis serves as a complement to the many insightful case studies presented in the existing literature to illuminate this link.

Our analysis is also related to the important work by Leemann and Mares (2011) and Mares (2015) on the politics of secret ballot in Imperial Germany. In that context, electoral corruption manifested itself mostly through intimidation of voters by employers and public officials. Mares (2015, chapter 7) shows that the demand for ballot reform (aimed at increasing the degree of secrecy in voting) in Prussia systematically came from politicians elected in constituencies with a high degree of economic diversification, presumably because electoral strategies based on intimidation were more costly in places with a more diversified employment structure. Our theory emphasizes how other aspects of modernization—income growth, urbanization, and education—increase the monetary cost of vote buying thereby reducing opposition to ballot reform. While the in-depth analysis of a single case, such as that of Prussia, has many advantages, the external validity will always be an issue. Our approach is to study the relationship between modernization and ballot reform (and the consequences thereof) using a

variety of research designs, ranging from cross country and cross state analyses to constituency level analysis within a given country. This, we believe, helps with external validity.

Section 2 presents the theoretical framework. Section 3 contains the results of the two event history studies. Section 4 examines the effects of the secret ballot on turnout. Section 5 discusses alternative mechanism and theories. Section 6 concludes. The supplementary material [intended for online publication] includes mathematical proofs and an evaluation of key assumptions, information on our coding of ballot reforms, definitions of all the variables used in the empirical investigation and their sources, and information on empirical robustness checks including instrumental variable estimations.

2 A Theory of Ballot Reform

We propose a new theory of ballot reform that formalizes the logic behind the three particular hypotheses that we test empirically.[4]

2.1 Assumptions

We consider a society with regular elections. The suffrage is not universal and voting is, initially, open. While we take the suffrage as given, the choice of ballot system is endogenous. In each period, two parties–party E and party R–compete in an election. Party E represents the old elite, while party R represents the (enfranchised) middle or working classes, henceforth called the radicals.[5] All N_E core supporters of party E can vote, while only some (N_R) of party R's core supporters are enfranchised. An (exogenous) increase in N_R represents a franchise extension. The old elite is outnumbered $N_R > N_E$. Voting is costly and some enfranchised voters may decide not to vote. The party allegiance of a voter is observable and the parties share policy preferences with their core voters. The party that gains the support of the majority of those who turn out to

[4] For alternative theoretical models of vote buying, see Dekel et al. (2008), Snyder (1991), Heckelman and Yates (2002), Dal Bo (2007), Baland and Robinson (2008) and Stokes et al. (2013).

[5] The parties should not be interpreted literally as particular historical political parties. Party E represents the old social elites whose claim to power is threatened by liberal democracy. Party R represents groups who stand to gain from liberal democracy. In the context of 19th century Britain, they can be exemplified by Radical Members of Parliament, the Chartists, or the Westminster Committee; in the context of 19th and early 20th century USA by the Mugwumps or Liberal Societies; and in the context of early 20th century Imperial Germany, by the liberal fraction of the parliament who proposed the use of standardized urns (Leemann and Mares, 2011, p. 7).

vote wins office. The historical context of our overall study does not suggest that the parties had the capacity (or incentive) to contest elections based on broad programmatic policy platforms. As in Keefer and Vlaicu (2008), the inability of parties to commit to programmatic policies is one of the factors that make vote buying attractive. Yet, voters would care, at least to a small extent, about which party is in power.[6] To capture this, we assume that the policy with party R–policy R–in power is better for voters of type R than the policy associated with party E–policy E–and vice versa. With this in mind, we can write the utility gain for a voter of type R or E of having "their" party in power as

$$\Delta_R \equiv u_R(R) - u_R(E) > 0 \quad (1)$$

$$\Delta_E \equiv u_E(E) - u_E(R) > 0, \quad (2)$$

where $u_i(j)$ is the utility that a voter of type $i \in \{E, R\}$ derives from policy $j \in \{E, R\}$. It is possible that the utility loss per capita to the elite associated with a switch in power from party E to party R is larger than the gain per capita to the radicals or vice versa. A neutral assumption is that gains and loses are of equal size and, as nothing substantial depends on it, we let $\Delta \equiv \Delta_R = \Delta_E$. We observe that Δ—the utilty the utility difference between the policy platforms of the two parties—is likely to increase over time as parties aquire the technolgies needed to commit to more programatic policy platforms.

The society has an infinitely long time horizon. Time is indexed by t and the common discount factor is $\beta \in (0,1)$. Since calender time plays no important role, we omit time index t when it is not strictly needed. There are two possible ballot regimes: open or secret ballot. Under secret ballot (SB), there is no electoral corruption; under open ballot (OB), votes can be bought and sold in a vote market.[7] Initially, the ballot is open, but the secret ballot may be adopted if the majority party proposes such a reform and the opposition party is unwilling to call a costly veto. Once secret ballot is introduced, it is not possible to go back to open ballot again.

Within a given period, five stages (A to E) evolve sequentially. What happens within each stage depends on the ballot regime.

[6] In some cases, this tension may manifest itself as a distributive conflict between the old elite and the majority of the electorate (Acemoglu and Robinson, 2006; Boix, 2003, chapter 1), but in other cases, the conflict may be about the type of public goods to provide, trade or education policy, or labor market regulation.

[7] Stokes (2005) along with the many detailed case studies in Kitschelt and Wilkinson (2007) show how party machines even in the presence of the secret ballot can use social networks to buy votes and how other forms of clientelism can survive. Accordingly, vote markets can exist even with secret voting, but they are clearly less effective. We make, for simplicity, the extreme assumption that the vote market shuts down with the arrival of the secret ballot, but it would be sufficient to assume that it is harder to buy votes under secret than under open ballot.

A. Planning. Under open ballot, the two parties (simultaneously) decide on how many opposition voters to target with a bribe in the upcoming election. We denote these targets by $n^b_{\sim jj}$ for $j \in \{E, R\}$.[8] The associated cost is $p_j v_{\sim j} n^b_{\sim jj}$ where p_j is the money offered to a voter of type j in exchange for his vote (to be determined below) and $v_{\sim j}$ is the marginal cost of raising funds for party $\sim j$. The parties care about policy, the cost of electoral bribery, and being in power. Their per-period expected payoffs are

$$V_E = f_E(u_E(E) + M) + (1 - f_E)u_E(R) - p_R v_E n^b_{ER} \quad (3)$$

$$V_R = (1 - f_E)(u_R(R) + M) + f_E u_R(E) - p_E v_R n^b_{RE}, \quad (4)$$

where f_E is the probability that party E wins the election and M is the utility value of political office. Under secret ballot, the parties do not, by assumption, buy votes.

B. Electoral turnout. The enfranchised voters decide whether they want to vote. Under open ballot, each voter compares the sum of his expressive benefit of voting (θ) and the expected utility value of the bribe (v^e_j) to the utility cost of voting ($c \in (0,1)$). He votes if $\theta + v^e_j \geq c$ for $j \in \{E, R\}$ where the expected utility value of the bribe is

$$v^e_j = \frac{n^b_{\sim jj}}{N_j} p_j \lambda_j \quad (5)$$

and λ_j is the marginal utility of income. Letting the expressive benefit θ be uniformly distributed on [0,1] for all voters, the number of voters of type j turning out to vote is

$$n^{OB}_j(n^b_{\sim jj}, N_j) = N_j(1 - c) + n^b_{\sim jj} p_j \lambda_j. \quad (6)$$

Turnout amongst voters of type j increases in the number of enfranchised voters of this type. More importantly, by offering bribes a party gives *opposition* voters an extra reason to turn out to vote. We notice the parties are not paying voters directly to turn out to vote. Rather this positive turnout effect is a side effect of vote buying. Under secret ballot, no bribes are offered and turnout is lower:

$$n^{SB}_j(0, N_j) = N_j(1 - c). \quad (7)$$

It is possible that parties, as it becomes impossibly to buy votes outright, start engaging in negative turnout buying, i.e., pay opposition voters to abstain. We do not model this possibility explicitly, but note that turnout would fall even further after the ballot becomes secret if this

[8] The notation $\sim j$ means "not j" , so if $j = E$, then $\sim j = R$ and vice versa.

were the case. This, in turn, reinforces our main results.

C. Vote buying. The voters who decided to vote show up at the polling station. Under open ballot, a vote market can operate. A voter of type j is willing to shift his allegiance to party $\sim j$ if offered at least his reservation price (p_j). The reservation price for a voter of type R is the monetary sum needed to compensate him for the economic loss of having policy E instead of policy R, and similarly for a voter of type E.[9] Formally, $p_j = \Delta\lambda_j^{-1}$ for $j \in \{E, R\}$, where we use the marginal utility of income (λ_j) to convert the utility differential (Δ) into a monetary amount. The radicals are poorer than the elite. Since the marginal utility of income falls with income ($\lambda_E < \lambda_R$), the reservation price is higher for voters of type E than for voters of type R. Under secret ballot, the vote market does not operate.

D. Polling. The election outcome depends on the relative electoral support of the two parties and on random events that might induce some voters to shift their allegiances (after bribes, if any, are paid). This induces a preference shock (η). The shock can, in principle, be positive (a shift to the radicals) or negative (a shift to the elite) and is introduced to capture the unpredictability of elections. Under open ballot, party E wins a majority amongst those who turn out to vote if

$$\frac{n_E^{OB} - \eta + g(v)}{n_E^{OB} + n_R^{OB}} \geq \frac{1}{2}, \tag{8}$$

where $v = \alpha_E n_{ER}^b - \alpha_R n_{RE}^b$ is the (productivity adjusted) difference between the number of votes bought by party E and party R. The function g, which is increasing, strictly concave and satisfies $g(0) = 0$, describes the exogenous vote buying technology. The two parameters (α_E and α_R) represent measures of vote buying productivity. As Baland and Robinson (2008) point out, the productivity of vote buying depends, among other factors, on the economic and social sanctions that the buyer can impose on the seller if the vote contract is broken. One of the ways in which modernization can reduce electoral corruption is, therefore, by undermining the scope for economic and social sanctions, i.e., by lowering α_j.

Under the assumption that η is distributed uniformly on the interval $[-\frac{\hat{\eta}}{2}, \frac{\hat{\eta}}{2}]$, the win probability of party E under open ballot is

[9] This is one way to conceptualize the cost of a vote. An isomorphic assumption is that a voter incurs an exogenous psychic cost voting against the party whose policy benefits him the most. We can then interpret Δ as this psychic utility cost. This would remove any instrumental element from the vote decision and have no consequence for the results of interest.

$$f_E(n_E^b, n_R^b, n_E^{OB}, n_R^{OB}) = \begin{Bmatrix} 0 \\ \frac{1}{\widehat{\eta}}\left(\frac{\widehat{\eta}}{2} + \frac{n_E^{OB} - n_R^{OB}}{2} + g(v)\right) \\ 1 \end{Bmatrix} \text{for} \begin{Bmatrix} \left(\frac{n_E^{OB} - n_R^{OB}}{2} + g(v)\right) < -\frac{\widehat{\eta}}{2} \\ \left(\frac{n_E^{OB} - n_R^{OB}}{2} + g(v)\right) \in -\frac{\widehat{\eta}}{2}, \frac{\widehat{\eta}}{2}] \\ \left(\frac{n_E^{OB} - n_R^{OB}}{2} + g(v)\right) > \frac{\widehat{\eta}}{2} \end{Bmatrix}. \quad (9)$$

The win probability of party R is $1 - f_E$. Under secret ballot, the vote market shuts down ($g(0) = 0$) and the win probability of party E is $f_E^{SB} \equiv f_E(0,0, n_E^{SB}, n_R^{SB})$. Since party R got more core supporters than party E, party E is likely to lose under secret ballot ($f_E^{SB} < \frac{1}{2}$). After the election, the winner implements its most-preferred policy and earns the office rent (M).

E. Ballot reform. Under open ballot, the winning party may propose that the secret ballot is adopted for future elections. While the majority party can implement its policy platform without the consent of the opposition, the opposition can, at a cost $\rho > 0$, veto such a proposal. If a reform proposal is made and not vetoed, then the new ballot regime applies for all future elections. If the ballot is already secret, it stays secret. The idea we want to capture with the veto is that reforming the ballot system is fundamentally different from day-to-day policy making and needs to pass a stricter test than the simple majority rule. A natural way to think of the veto is that it represents the ease with which an upper chamber can block proposals from the lower chamber and at what cost.[10] For example, even in the 19th century where the British House of Lords had real power, most day-to-day policy making was a matter for the House of Commons, but institutional reforms would be subject to a veto constraint from the Lords, as the politics of the Great Reform Act of 1832 illustrates. The veto cost captures this reality is a simple and workable way.

We analyze the model in two steps. First, we characterize the equilibrium in the vote market. Second, we study the reform process.

2.2 The vote market

If the secret ballot was introduced in the past, party E wins with probability $f_E^{SB} < \frac{1}{2}$. If

[10] Aidt and Giovannoni (2011) develop a theory of constitutional rules that rationalizes the distinction between special and day-to-day rules.

not, the vote market can flourish and the two parties must, in stage A, decide how many opposition voters to target with bribes in the upcoming election, anticipating events as they unfold in the subsequent stages. Using equations (3) and (4) and recalling from equation (6) that $\frac{\partial n_j^{OB}}{\partial n_{\sim jj}^{b}} = \Delta$ and $p_j = \Delta\lambda_j^{-1}$ for $j \in \{E, R\}$, the two first order conditions governing these choices can be stated as follows:

$$n_{ER}^{b}: \frac{1}{\hat{\eta}}\frac{\partial g}{\partial v}(v) \leq \frac{\upsilon_E \Delta}{\lambda_R \alpha_E (\Delta + M)} + \frac{\Delta}{2\hat{\eta}\alpha_E} \text{ with} = \text{if } n_{ER}^{b} > 0 \quad (10)$$

$$n_{RE}^{b}: \frac{1}{\hat{\eta}}\frac{\partial g}{\partial v}(v) \leq \frac{\upsilon_R \Delta}{\lambda_E \alpha_R (\Delta + M)} + \frac{\Delta}{2\hat{\eta}\alpha_R} \text{ with} = \text{if } n_{RE}^{b} > 0. \quad (11)$$

The two terms on the right-hand side shows the cost of buying an extra vote. The first term is the direct (utility) cost per vote. The second term represents the reduction in the win probability of the party that offers the extra bribe because bribery increases turnout amongst opposition voters. The left-hand side represents the increase in the win probability of allocating an extra dollar to buying opposition votes before adjusting for productivity differences (captured by α_j) and is the *same* for the two parties. An implication, then, is that only the party with the lowest productivity adjusted cost buys votes.[11] Under Assumption 1, this party is the minority party representing the elite.

Assumption 1 $\frac{\upsilon_E}{\lambda_R \alpha_E (\Delta + M)} + \frac{1}{2\hat{\eta}\alpha_E} < \frac{\upsilon_R}{\lambda_E \alpha_R (\Delta + M)} + \frac{1}{2\hat{\eta}\alpha_R}$.

Assumption 1 is likely to be satisfied because voters of type R are cheaper to buy than voters of type E, as they are poorer ($\lambda_R > \lambda_E$); because the elite can raise funds at lower cost than the radicals ($\upsilon_E < \upsilon_R$); and because the elite is more effective at buying votes than the radicals ($\alpha_E > \alpha_R$), as they can use social sanctions more effectively. The equilibrium win probability of party E is

$$f_E^{OB} \equiv f_E(n_{ER}^{b^*}, 0, n_E^{OB}(0), n_R^{OB}(n_{ER}^{b^*})) > f_E^{SB}, \quad (12)$$

where $n_{ER}^{b^*}$ represents how many radical voters party E offers bribes to at equilibrium.

The equilibrium in the vote market ($n_{ER}^{b^*}$) and the win probability of the minority party, f_E^{OB}, are affected by three key parameters, α_E, λ_R and N_R. Firstly, an increase in the transaction

[11] See the supplementary material appendix S1 for proofs and derivations.

cost (a fall in α_E) reduces the marginal benefit of each bribe and, as a consequence, party E targets fewer opposition voters and the vote market shrinks. This gives radical voters one less reason to vote and their turnout falls. Yet, the net effect is that party R is more likely to win a majority. Secondly, a fall in the private marginal value of income (λ_R) increases the price of a vote. Intuitively, radical voters value money less and so require a higher monetary compensation to shift their allegiance. This increases the cost of bribery and party E buys fewer votes. The vote market shrinks, and although turnout falls amongst voters of type R, the consequence is, again, that the win probability of party E goes down. An exogenous franchise extension–an increase in N_R–does not affect the absolute number of opposition voters party E wants to bribe but the relative share drops and, as a consequence, the win probability of party E falls. Suffrage reform makes vote buying uneconomical.

We contend that "modernization"–income growth, urbanization and education–are systematically related to the transaction cost of vote buying and to the marginal value of income. The latter relationship is immediate: income growth combined with the standard assumption of diminishing marginal utility of consumption implies that the private marginal value of income (λ_R) decreases with income growth. This increases the price of a vote and gradually reduces the volume of transactions in the market for votes. The first relationship between "modernization" and the transaction cost of vote buying (α_E) is more complex and requires elaboration. First, a vote market operates most effectively in environments with a high degree of economic dependency, social control and deference to authority. The examples of Britain before the Ballot Act of 1872 and Imperial Germany, which nominally had secret ballot from 1871, demonstrate how the social structures of small towns and villages enabled local authorities and elites to control elections through outright bribery or through social deference.[12] Industrialization and urbanization disturb this equilibrium by opening up new economic possibilities for working and middle class voters and by making them economically mobile. This destroys traditional patron-client networks, which, once they are gone, are difficult to re-create (Hicken, 2007). Anderson (1993, p. 1458) paraphrases the argument about urbanization in her analysis of Imperial Germany as follows: "supported by the anonymity of the new megalopolis, [...], voters escape both the meshes of deference and the terrors of officials and employers. Town air makes free. Such cities were simply too large for traditional authorities to control." The transition from a

[12] Anderson (1993), Lee (1994, p. 144) and Mares (2015).

static agrarian economy to a dynamic industrial economy with deeper markets and economic specialization, therefore, makes it harder for the old elite to enforce and monitor vote contracts, and the transaction cost of vote buying shoots up. Moreover, economic development "weakens the power of the landlord class and strengthens subordinate classes. The working and the middle classes [...] gain an unprecedented capacity for self-organization due to such developments as urbanization, factory production, and new forms of communication and transportation" (Rueschemeyer et al., 1993, p. 75). By allowing new political parties and special interest groups (e.g., unions) to emerge, this also serves to undermine old structures of deference. While it is clear, then, that urbanization breaks down structures that previously enabled vote markets, large urban electorates are not a guarantee that electoral corruption is eliminated. The experience of many American cities, e.g., New York, in the Gilded Age clearly demonstrates that politicians and parties determined to control the electorate can find new ways of doing so (Fredman, 1968). Urbanization may, therefore, have encouraged electoral corruption in some contexts (e.g., in US migration cities), while having discouraged it in others (e.g., in Britain and Germany). Second, modernization is associated with a rise in literacy and since "education presumably broadens men's outlooks, enables them to understand the need for norms of tolerance, restrains them from adhering to extremist and monistic doctrines, and increases their capacity to make rational electoral choices" (Lipset, 1959, p. 79), this also makes it harder to enforce vote contracts.

Taken together, these arguments suggest that modernization, either directly through the effect on the marginal value of income or indirectly through urbanization and the rise of education, reduces the incentive of parties, which benefit from electoral corruption, to use vote buying as an electoral strategy. Consequently, as modernization progresses under open ballot, vote markets continue to exist but they become less and less important and fewer and fewer votes are bought and sold in them. This, in turn, has, as we shall now show formally, implications for the likelihood of ballot reform.

2.3 Ballot reform

At stage E, the winning party can, if the ballot regime at time t is open ballot, propose to adopt the secret ballot for future elections. Clearly, party E has a no reason to do so. A reform proposal must come from party R which stands to improve its electoral prospect ($f_E^{OB} > f_E^{SB}$).

The implication, then, is that a reform proposal is tabled at time t only if party R wins the election under the open ballot. This happens with probability $1 - f_E^{OB}$. As discussed above, both modernization and franchise extension increase the cost of buying votes in the market. Party E, therefore, rationally buys fewer votes and its electoral prospect deteriorates. This increases the likelihood that ballot reform is proposed. Whether it will be reformed, however, depends on the elite's willingness to veto.

Suppose that party R wins the election at time t and proposes a reform. Party E can veto this proposal at cost $\rho > 0$. By doing so, the ballot remains open in period $t + 1$, but another veto may have to be called if party R happens to win again and so on. To evaluate the elite's willingness to veto, we compare the present discounted value of a veto to the present discounted value of accepting the ballot reform:

$$W_E(\text{no veto}) = \beta \frac{(1-f_E^{SB})u_E(R)+f_E^{SB}(u_E(E)+M)}{1-\beta} \tag{13}$$

$$W_E(\text{veto}) = \beta \frac{(1-f_E^{OB})u_E(R)+f_E^{OB}(u_E(E)+M)}{1-\beta} \tag{14}$$

$$- \frac{\beta v_E \Delta \lambda_R^{-1} n_{ER}^{b*}(.)}{1-\beta} - \frac{\beta(1-f_E^{OB})\rho}{1-\beta} - \rho.$$

Under secret ballot, power simply alternates between the two parties according to the win probability f_E^{SB}, which is determined by the relative turnouts of the two groups and random events. This is what the numerator of equation (13) represents. Since not calling a veto at time t means a permanent switch to secret ballot starting at time $t + 1$, this expected value is discounted by the factor $\frac{\beta}{1-\beta}$. A veto preserves the open ballot for another period. Under open ballot, power alternates between the two parties according to the win probability f_E^{OB}, which takes into account the vote market. This is represented by the first term of equation (14). On top of this comes the cost of buying votes each period (the second term) and the veto cost, which is paid after each electoral defeat in the future (the third term). The fourth term is the veto cost at time t.

The elite's willingness to veto is the difference between $W_E(\text{veto})$ and $W_E(\text{no veto})$. We can state the condition under which the elite will veto as:

$$\frac{\beta}{1-\beta}\{(f_E^{OB} - f_E^{SB})(\Delta + M) - v_E \Delta \lambda_R^{-1} n_{ER}^{b*}(.)\} \geq \rho + \frac{\beta(1-f_E^{OB})\rho}{1-\beta}. \tag{15}$$

The left-hand side represents the net benefit of the veto. This depends positively on the electoral

advantage that the vote market gives party E ($f_E^{OB} - f_E^{SB}$) and negatively on the cost of bribery. The right-hand side represents the veto cost.

The costs and benefits of a veto depend on the two modernization parameters (α_E and λ_R) and how many of party R's core voters have been enfranchised (N_R). Modernization–represented either by an increase in the transaction cost or by income growth–makes the elite less likely to veto. The reason is that modernization reduces the electoral advantage gained from the vote market (the gap between f_E^{OB} and f_E^{SB} narrows) while at the same time the frequency with which a veto is required goes up because party R is more likely to win office. Franchise extension also makes the elite less likely to veto a subsequent ballot reform. Since the total number of opposition voters targeted is unaffected, the gap between f_E^{OB} and f_E^{SB} is unaffected by an increase in N_R. However, party R becomes more likely to win elections and, for that reason, the expected veto cost increases and this is what makes party E less willing to call a veto.

We summarize these insights in three hypotheses which we test empirically.

Hypothesis 1 *(The modernization hypothesis). Modernization increases the likelihood that the secret ballot is adopted.*

Hypothesis 2 *(The franchise hypothesis). An extension of the franchise increases the likelihood that the secret ballot is adopted subsequently.*

Hypothesis 3 *(The turnout interaction hypothesis). The secret ballot is associated with a fall in electoral turnout and this fall is smaller where modernization has progressed the most.*

The third hypothesis requires elaboration. Under open ballot, modernization gradually reduces turnout. This is because as the price of a vote goes up, party E becomes less willing to buy votes. As a consequence, from the point of view of a voter of type R, the value of voting and, in expectation, be offered a bribe falls and more and more such voters decide to stay at home. Our theory does not necessarily suggest that the vote market has disappeared by the time the secret ballot is introduced, only that fewer votes are being bought and sold than in previous times. Accordingly, secret ballot also leads to a fall in turn out because at that point, what remains of the vote market shuts down altogether. To see how the two effects interact to generate

hypothesis 3, imagine two societies which happen to introduce the secret ballot at the same time. The only difference between the two is that one is, say, richer than the other. Consequently, the vote market is less vibrant and fewer voters bother to vote in the rich than in the poor country in the years leading up to the reform. When the reform happens, the fall in electoral turnout is smaller in the rich than in the poor country.

Our theory is developed in the context of direct vote buying. In reality indirect vote buying through employment relationships or other favours played an equally important role in many countries, most notably perhaps Imperial Germany (Ziblatt, 2009; Leemann and Mares, 2011; Mares, 2015) or Chile (Baland and Robinson, 2008). So, do our predictions regarding "modernization" and the secret ballot also apply to a market dominated by indirect vote buying? The answer is that they do. To see this, notice that we can reinterpret the price (p_j) as an effort cost incurred by the parties. Effort can coerce voters to "accept" the utility cost Δ, which we might reinterpret as a psychic cost of voting against once preferred party, because the sanction associated with not doing so is greater. "Modernization" reduces the scope for social control by increasing the effort cost required to buy votes and in that way undermines the value of indirect vote buying.[13]

3 Event history studies of the secret ballot

The aim of our event history study is to explain the timing of the secret ballot. We model the (conditional) probability that a country or a state which has not yet adopted the secret ballot adopts it in a given year as a function of quantitative measures of modernization, the franchise, and other potential determinants of ballot reform.

We explore two different historical samples—Western Europe plus off-shoots (1820-1913) and the US states (1840-1950)—that cover the relevant period during the 19th and 20th centuries when the secret ballot replaced the open ballot. We are aware that this is a small subset of the countries or states that have undergone the transition from open to secret ballot. Given the constraints imposed by the availability of quantitative historical data on modernization, in

[13] The model is based on a number of simplifying assumptions which we believe make sense in the context. We have, however, investigated the robustness of the results to many alternative assumptions. The details of these permutations are included in the supplementary material appendix S1. The bottom line is that the predictions are robust.

particular, GDP per capita, it is not possible to expand the Western Europe plus off-shoots sample further, say, with countries from Eastern Europe.[14]

We are interested in the year in which the secret ballot became effective in a country or a state. By effective we mean that the ballot rules were such that voters could cast their vote in private (i.e., without the vote choice being observed and unduly influenced by others) and that it was impossible ex post to reconstruct how a particular voter voted. We focus on when the secret ballot became effective, because this is what our theory suggests is relevant. In the theory, the secret ballot is introduced when the benefits to the elites are less than the veto cost. Introducing secret ballot de jure but with ways to circumvent secrecy works similarly to the veto in our theoretical model. From a practical point of view, de jure secrecy may not be de facto secrecy. This may be so because of non-standard ballot paper of different collors or shapes; because of non-standard urns (for example, a small urn which preserves the order of voting can reveal the choices of individual voters ex post); or because of the absence of a place where the ballot can be filled in and placed in the urn without outsiders being able to observe the act (Mares, 2015).

A number of different arrangements can in principle and in practise make the ballot secret in this sense. One such arrangement is the Australian ballot. The Australian ballot requires, at least, two things, First, all parties/candidatesare listed/printed on the same ballot paper, and, second, these ballot papers are printed by the state at the public's expense and distributed only at the polling stations and placed in a standard-sized urn in privacy. Another arrangement is the so-called "ballot and envelope" system, practiced in, e.g., France, Sweden and Spain. It requires the state to print the ballot papers, but one ballot paper is printed per party or candidate. The voter then chooses the paper for the party/candidate he wants to cast his vote for, puts it in an envelope and deliver it to the ballot box. This type of arrangement (and others like it) is more open to abuse than the Australian ballot but de jure makes the ballot secret. In coding the year in which secret ballot was introduced in the countries covered by the Western European plus off-spring sample, our rule is that a new ballot arrangement qualifies as "secret ballot" if our sources indicate that the relevant change in balloting procedures effectively made the act of voting private. In most cases, it is straight forward to establish the year of adoption thus defined, but in some cases, most notably Germany, there is disagreement about when the ballot became

[14]The problem with extending the sample to Eastern Europe is that reliable historical data on GDP per capita and other aspects of modernization are very limited. The evolution of democratic institutions in many eastern European countries during the 19th century, however, followed a pattern not all that dissimilar to that in Western Europe. In particular, Aidt and Jensen (2014) report that the dynamics of suffrage reform was comparable in the two parts of Europe. We do, therefore, not believe that sample selection bias is a major issue, although we cannot rule it out altogether.

effectively secret. We use 1913, when standard urns were introduced in the main specification, but note that our results are unchanged if we instead use 1903 when ballot envelopes and isolating places where introduced (Mares, 2015). In the supplementary material (appendix S2), we provide information on how we coded the year of adoption for each country and why. For the sample of US states, we code the year in which the ballot in a state became Australian in state-wide elections based on the information provided by Lott and Kenny (1999).

The dependent variable $reform_{it}$ is coded one if country (or state) i introduced the secret ballot in year t and as zero in the years before and after that. A country (or state) drops out of the sample in the year after its adoption. We assume that all polities were at "risk" of adopting the secret ballot from the beginning of the (relevant) sample period or when they became independent. This assumption is justified by historical facts. By the late 18th century the secret ballot was widely debated in liberal circles and supported by philosophers, such as Jeremy Bentham. In Britain, the secret ballot was proposed by the Westminster Committee as early as 1780 (Schofield, 2004). Thus, the principle of secret ballot was known and on the agenda in many countries by 1820. It is, therefore, reasonable to assume that the countries (states) under consideration could have adopted the secret ballot from 1820 (1840) onwards.

As in Beck et al. (1998), we estimate a discrete logistic model:[15]

$$P(\text{reform}_{it} = 1 | x_{it}, m_{it-1} = 0) = \frac{1}{1+e^{-(x_{it}\gamma + H(.))}}. \quad (16)$$

The variable m_{it-1} is an indicator variable equal to zero in each year before introduction of the secret ballot and equal to one thereafter. Through the function $H(.)$, we allow the hazard rate to be time-varying and a function of the number of years a country has been at risk of adoption.[16] This is important because the likelihood of adoption may increase as time passes. By taking duration dependence into account, we avoid confounding any modernization or franchise effect with a spurious correlation between two upwards trending variables. The vector x_{it} represents three main groups of explanatory variables. The first group contains indicators of modernization, such as the log of real GDP per capita, the urbanization rate, and measures of education attainment standards. The second group contains variables related to the extension of the

[15] Beck et al. (1998) show how the logit model can be derived from the Cox proportional hazard model. The logit formulation has the advantage of being easier to interpret, of allowing for flexible estimation of duration dependence, and of making it possible to adjust for rare events. For these reasons, it is widely used in event history studies.

[16] An alternative to this is to use calendar time to track the baseline hazard. This would, for example, make sense if the trend in the hazard rate is global. We have re-estimated all the logit models using a common time trend defined on calendar time and it makes no difference to the results. We prefer the specification with clock time since it allows for country-specific idiosyncrasies.

franchise. The third group contains variables that capture alternative causes of ballot reform. Many scholars emphasize the importance of landholding and income inequality in relation to democratization in general. We expect that landholding inequality makes the secret ballot less likely because inequality along this dimension often reflects a social order that favors vote buying. Other scholars, e.g., Wejnert (2005), Gleditsch and Ward (2006), and Leeson and Dean (2009), emphasize international diffusion of democracy. Governments in one polity may learn from political reforms—in our case ballot reforms—in other polities, and more so from those which are either linguistically or physically nearer. Finally, we control for population size.

Table 1: The effect of modernization of the probability of secret ballot in Western Europe plus off-shoots 1820-1913.
Dependent variable: *reform*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log(real *GDP* per capita)	6.36**		6.62*		5.43**		5.32		0.15*
	[2.44]		[1.77]		[2.11]		[1.44]		[1.99]
Urbanization rate		0.01***	-0.001			0.09***	-0.001		
		[3.65]	[-0.21]			[3.38]	[-0.01]		
First principal component				1.46***				1.29***	
				[3.27]				[2.94]	
Electorate/adult population	0.037	0.055**	0.034	0.05*	0.034	0.051**	0.033	0.033	0.002
	[1.26]	[2.30]	[0.94]	[1.87]	[1.19]	[2.17]	[0.94]	[0.94]	[0.92]
Gini coefficient	28.49	12.33	28.35	16.80	25.36	12.03	26.21	26.21	-1.041
	[1.44]	[0.66]	[1.39]	[0.93]	[1.30]	[0.66]	[1.30]	[1.30]	[-1.12]
log(population)	-0.50**	-0.76***	-0.46	-0.71**	-0.41*	-0.66**	-0.39	-0.39	0.13
	[-2.04]	[-2.80]	[-1.35]	[-2.54]	[-1.72]	[-2.47]	[-1.18]	[-1.18]	[1.65]
Learning	-1.18	1.2	-1.26	0.22	-1.03	0.93	-1.03	-1.03	0.001
	[-0.94]	[1.56]	[-0.95]	[0.27]	[0.84]	[1.44]	[-0.78]	[-0.78]	[0.03]
Joint significance of modernization variables			**6.24****				**4.91***		
Control for duration dependence	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	779	732	732	732	779	732	732	732	779
Estimation method	Logit	Logit	Logit	Logit	Rare Events	Rare Events	Rare events	Rare events	Fixed Effects OLS

Notes: Robust z-statistics correcting for clustering by country in brackets (i.e., we allow the errors to be correlated over time within countries, but not between countries). Constants not reported. All estimations based on sample of 14 countries. The small number of countries raises the concern that the standard errors corrected for clustering are misleading. We have done two checks to judge the relevance of this issue. First, we use more conservative critical values for the t-distribution based on G-K degrees of freedom with G being the number of clusters and K the number of variables (Angrist and Pischke, 2009). Second, we apply a wild cluster bootstrap in the linear probability model (Cameron et al., 2008). None of these techniques indicate that the inference based on the clustered standard errors reported in the table are misleading. *** p<0.01, ** p<0.05, * p<0.1.

3.1 Western Europe plus off-shoots

The Western Europe plus off-shoots sample covers, for the period from 1820 to 1913, 11 Western European countries[17] plus the USA, Canada, and New Zealand. The first country in the sample to introduce the secret ballot was the Netherlands in 1849; the last ones were France and Germany in 1913. Before the secret ballot, electoral corruption was widespread. In many Western European countries vote buying was, during the 19th century, concentrated in the countryside where social control and employment relations made it relatively easy for the landed elites to run vote markets (e.g., Ziblatt, 2009; Seymour, 1915, p. 433; Mackie, 2000). In the USA, vote markets were also widespread in the countryside, but they were often particularly vibrant in the big migration cities where party machines exploited that colored voting papers, indicating party choice, could be handed out at the polling stations (Cox and Kousser, 1981; Ostrogorski, 1964).

In Table 1, we report the results of the event history study. We see that the coefficients on the two modernization variables, *urbanization rate* and *real GDP per capita*, are positive and statistically significant (columns one and two).[18] When they enter together (column three), they are jointly significant and *real GDP per capita* remains significant on its own. The two variables are positively correlated, with a correlation coefficient of 0.75, and are likely to capture the same underlying concept of modernization. It, therefore, makes sense to extract the principal components. In column four, we replace the two modernization variables with the first principal component which correlates positively with *real GDP per capita* and with *urbanization rate*.[19] This component has a positive and statistically significant coefficient. Ballot reform is a rare event and this may bias the estimates. King and Zeng (2001) propose a logit estimator that deals with this. In columns five to eight, we report results using this estimator. The two modernization variables remain individually significant and jointly significant when entered together or as the first principal component. To quantitatively evaluate the modernization effect, we may begin by noticing that, based on the specification in column three, the probability of a ballot reform in a given year conditional on the average values of the covariates is 0.3 percent. Suppose now that *real GDP per capita* increases by one standard deviation, keeping all other variables at their average values. As a

[17] The Netherlands, Switzerland, United Kingdom, Belgium, Norway, Denmark, Finland, Austria, Sweden, France and Germany.
[18] Systematic information on enrollment in education or other proxies for education levels are not available for a sufficiently long time span for this sample. As a rough proxy, we have coded a dummy variable which takes the value of zero for each country till enrollment in primary education amongst 7 to 14 year old reaches 60 percent and then is one thereafter. The 60 percent cut off is the lowest cut off we can capture with the spare data available. We note that this variable is not significant most likely due to measurement error [not reported].
[19] The factor loadings on the first principal component are equal to 0.93 for both variables.

consequence of this the predicted probability of reform in a given year increases to 2.6 percent. While ballot reform remains an unlikely event in any given year, we observe that the likelihood increases many fold as GDP per capita increases.[20]

We use information on the number of eligible voters as a proportion of the adult population, *electorate/adult population*, to test the franchise hypothesis. The results reported in Table 1 are not favorable. Although the point estimate is positive, it is only statistically significant in specifications where *real GDP per capita* is excluded. Overall, it was not the pre-secret ballot expansion of the suffrage that triggered the secret ballot.

Income inequality, as measured by the *gini coefficient*, is (statistically) unrelated to the timing of the secret ballot. This runs counter to other recent evidence on the effect of inequality on democratization.[21] The control variable *population* always has a negative coefficient, but is usually not significant, suggesting that scale effects were unimportant. The variable *learning* captures diffusion effects. It is a "distance" weighted index of reforms in neighboring countries, where we use the information on linguistic similarities provided by Fearon (2003) to measure distance. Despite the fact that the adoptions of the secret ballot cluster in the 1870s, we find little evidence that social learning was important. A similar result is reported in Aidt and Jensen (2014) in relation to international diffusion of suffrage reforms. These results contradict the strong evidence on international diffusion of democracy coming from studies of large panels of countries, either in the second half of the 20th century (e.g., Gleditsch and Ward, 2006) or the entire period from 1820 or 1850 to 2000 (Persson and Tabellini, 2009 or Leeson and Dean, 2009). These studies focus on composite democracy measures (mostly the Polity IV index). Our results pertain to the first wave of democratization and to *specific* aspects of the package of democratic institutions. It appears that international diffusion of specific institutions was not all that important for the early democratizations in Western Europe.

The results reported in Table 1 are estimated from a combination of between and within country variation in modernization, the size of the electorate, inequality, etc. It is, therefore, possible that the correlations are driven by the same unobserved factors and that they are coincident rather than causal. To show that this is most likely not the case, we have undertaken two additional tests, both of which force us to work with a linear probability

[20] We have investigated the possibility of a nonlinear relationship between modernization and the secret ballot. For the Western Europe plus off-shoots sample, we find evidence of an inverted U-shaped relationship between *real GDP per capita* and the adoption probability, but this fails to replicate in the US states sample.

[21] Ziblatt (2008), Boix (2003, chapter 2), or Ansell and Samuels (2010).

model.[22] Firstly, we have estimated a linear fixed effects model which takes into account country-specific unobserved heterogeneity. We report the results in column nine of Table 1. We observe that the coefficient on the main modernization variable, *real GDP per capita*, now estimated only from variation within each of the 14 countries over time, is significant with a robust t-statistic of 1.99. Secondly, we have estimated the linear probability model with an instrumental variables estimator. We use a weighted index of the level of (log) real GDP in other countries as an instrument for (log) GDP per capita in a particular country. This is similar to Acemoglu et al. (2008), except that we use inverse distance rather than trade shares as weights. The main threat to the validity of the instrument is the possibility that an increase in real GDP per capita abroad induces ballot reform and this experience diffuses internationally. We deal with this by including the variable *learning*, which controls for the diffusion process, in the estimations. We use a measure of revolutionary shocks in neighboring countries as an instrument for the extension of the franchise in particular countries (see Aidt and Jensen (2014)). The instrumental variables results are reported and discussed in appendix S4 (Tabe S4) in the supplementary material. They show that the baseline results are robust and may be given a causal interpretation.

3.2 US States

The first US state to adopt the secret ballot, which for this sample we define as Australian ballot**,** statewide was Massachussts in 1888 and the last was South Carolina in 1950. Before the Australian ballot, electoral corruption was widespread. Bensel (2004) notes that the introduction in many states of the ticket system prior to the Australian ballot provided some privacy for voters, but also stresses that the degree of secrecy should not be overstated. The issue was that parties were able to ascertain voting behavior because they supplied voters with ballot papers (Lehoucq, 2007). Fredman (1968, p. 22) describes how it worked: "the simplest form of bribery occurred when ballot peddlers or district captains paid a voter as he emerged from the polling place. To check that he actually used the ballot it was colored or otherwise recognizable and the compliant voter was followed up to the booth." McCook (1892) estimates that sixteen percent of voters in Connecticut were up for sale at prices ranging from two to twenty dollars. Stokes et al. (2013) emphasize that the Australian ballot

[22] The linear probability model is only a good approximation, if events in the tails are unimportant. Since adoption of the secret ballot is an unlikely event, the tails are likely to be important, and we are careful not to read too much into the size of the point estimates obtained with the linear probability model.

reduced the effectiveness of vote buying in the US, and that it did so by diminishing "the observability of voter's choices" (p. 183). The most corrupt 19th century state elections are said to have occurred in New York and San Francisco. The reason was the high concentration of poor voters and recent immigrants unused to the franchise (Fredman, 1968, p. 25). Congleton (2011, p. 560) notes that the introduction of the Australian ballot "allowed votes to be cast without fear of rebuke by landlords or employers". Taken together this suggests that the ballot became *effectively* secret in the US states with the arrival of the Austrialien ballot.

We use three quantitative indicators of modernization–*real income per worker*, *average years of schooling*, and *urbanization rate*. From Table 2, we see that each of the modernization variables is positively and significantly correlated with the timing of the secret ballot (columns one to three). When included together in column four, they are jointly significant and *real income per worker* and *average years of schooling* remain significant on their own. We replace the three variables with the first principal component in column five and note it is highly significant.[23] These results are robust to alternative estimation techniques, as shown in columns five and six.

To evaluate the size of the modernization effect, suppose that *real income per worker* increases by one standard deviation. Using the specification in column four and keeping all other variables at their average values, this increases the predicted probability of a ballot reform from 0.21 to 0.37 percent. This effect is smaller than in the Western Europe plus off-shoots sample, but still substantial.

Unlike in Western Europe, 60-90 percent of adult (white) males was enfranchised already in the 1840s. Nevertheless, the states applied various strategies to *de facto* restrict the suffrage. These included requiring payment in full of poll taxes and subjecting potential voters to literacy tests. These steps served to keep poor and illiterate males off the election roll, often aimed at disenfranchising African Americans (Keyssar, 2009, chapters 4 and 5). Women's suffrage rights also varied and some of the frontier states granted women the right to vote long before it became mandatory in 1920 (Lott and Kenny, 1999; Braun and Kvanicka, 2013). We use these restrictions to capture over-time and across-state variation in the size of the electorate. From Table 2, it is clear, however, that these restrictions had very little impact on the secret ballot. We must, again, conclude that there is little evidence supporting the franchise hypothesis.

We use the share of land held by the 20 percent largest farms (*land inequality*) to

[23]The factor loading are 0.9291 (for income), 0.9063 (for schooling) and 0.8703 (for urbanization).

capture landholding inequality (Galor et al., 2009). Information on this share is only available from 1880, so we lose over half the observations when we include it in the specification shown in column seven of Table 2. Yet, landholding inequality is highly significant and exerts a negative impact on the secret ballot. This is in line with the findings of Ziblatt (2008) and Ansell and Samuels (2010; 2015) in other contexts. We find little evidence that scale effects or social learning mattered amongst the US states.

In appendix S4 in the supplementary material, we discuss a number of additional robustness checks. This includes instrumental variable estimation based on the same strategy as for the Western Europe plus off-shoots sample. The results are consistent with a causal interpretation of the relationship between modernization and secret ballot. Kuo and Teorell (2013) report independent evidence which further supports such an interpretation and suggests that the Australian ballot reduced vote buying and intimidation in US state elections between 1860 and 1930. Moreover, the case study of Stokes et al. (2013, chapter 8) also corroborates the hypothesis that modernization helped end clientelism in the USA.

Table 2: The effect of modernization of the probability of secret ballot in US states, 1840-1950. Dependent variable: *reform*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(real income per worker)	1.488***			1.239**		1.194**	0.601
	[3.616]			[2.374]		[2.300]	[1.016]
Average years of schooling		0.525***		0.324*		0.311*	0.784***
		[3.734]		[1.910]		[1.843]	[3.047]
Urbanization rate			0.0218**	-0.00661		-0.00492	0.0175
			[2.046]	[-0.384]		[-0.288]	[0.966]
First principal component					0.581***		
					[2.985]		
Women's suffrage	0.205	-0.122	0.106	0.156	0.195	0.187	-0.168
	[0.455]	[-0.264]	[0.215]	[0.331]	[0.426]	[0.399]	[-0.278]
Literacy test	-0.84	-0.876	-0.864	-0.775	-0.849	-0.717	-0.502
	[-1.212]	[-1.191]	[-1.327]	[-1.086]	[-1.276]	[-1.011]	[-0.934]
Poll taxes	-0.307	-0.45	-0.57	-0.254	-0.335	-0.203	-0.324
	[-0.817]	[-0.911]	[-1.358]	[-0.614]	[-0.881]	[-0.495]	[-0.637]
Land inequality							-31.1***
							[-3.670]
log(population)	-0.068	-0.28***	-0.201	-0.0877	-0.118	-0.104	-0.109
	[-0.651]	[-2.620]	[-1.481]	[-1.027]	[-1.247]	[-1.228]	[-0.923]
Learning	-0.055	-4.292	-12.19	-5.353	-9.686	-5.582	9.758
	[-0.005]	[-0.363]	[-0.941]	[-0.385]	[-0.771]	[-0.404]	[0.813]
Joint significance of the modernization variables				20.7***		20.1***	18.1***
Joint significance of the suffrage variables	2.33	2.02	3.40	1.72	2.70	1.57	1.49
Observations	2,239	2,214	2,251	2,186	2,186	2,186	738
Number of states	45[a]	48	47[b]	44	44	44	44
Estimation method	Logit	Logit	Logit	Logit	Logit	Rare events	Logit

Notes: Robust z-statistics correcting for clustering by state in brackets; a. Income data are missing for South and North Dakota and Oklahoma; b. Urbanization is missing for Idaho. All estimations control for duration dependence. *** p<0.01, ** p<0.05, * p<0.1.

4 Turnout and vote buying

The evidence presented so far supports the view that modernization systematically affects the timing of the secret ballot. The underlying mechanism, however, remains unclear. In our theory, the causal link runs through the vote market. The available quantitative evidence on the operation of vote markets comes from a handful of insightful studies of particular markets.[24] We provide new evidence on the interaction between modernization and the secret ballot by testing the turnout interaction hypothesis. The logic behind this hypothesis is that modernization gradually reduces the volume of trade in the vote markets operating under open ballot and that the secret ballot, in one go, elimintes (most of) what remains of these markets. As argued, e.g., by Converse (1974), Rusk (1974) and Heckelman (1995), this reduces electoral turnout. The elimination of the vote market is, however, not the only possible explanation for a fall in turnout after secret ballot. For example, it is possible that there is a shift from vote buying towards negative turnout buying (i.e., to pay expected opposition voters to stay home) after a reform. The seminal study by Cox and Kousser (1981, p. 656) of newspaper reports in New York State about instances of electoral corruption before and after the introduction of the secret ballot in 1890 provides examples of this. It is also possible that turnout will increase as a consequence of the secret ballot. This would be the case if the expressive benefit of voting shoots up by making the act of voting private or if the secret ballot "protects" voters and gives them the freedom to vote how they like. A possible example of this is Imperial Germany where the adoption of the secret ballot was correlated with an increase in turnout. Turnout could also increase if parties start using "brokers" to deliver blocks of votes (Stokes et al., 2013). The idea is that party leaders can contract with those brokers, who themselves use personal connections and knowledge of individuals to foster favorable turnout. This leads to positive turnout buying and if this is taking place at a sufficiently large scale, then turnout would not fall as a consequence of the secret ballot. We notice, however, that all these additional effects (which are not captured by our model) operate ex post, i.e., after the ballot has become secret. Our theory and the turnout interaction hypothesis is about what happens to the vote market and turnout ex ante, i.e., while the ballot is still open. What we argue is that the fall in turnout, documented by Heckelman (1995), is

[24] Ziblatt (2009) studies electoral corruption in Imperial Germany between 1871 and 1912; Cox and Kousser (1981) study voting buying in rural areas of New York state in 1890s; Baland and Robinson (2008) study the vote market in Chile in the 1950s; Stokes (2005) studies vote buying in Argentina in 2002; and Collier and Vicente (2012) study electoral corruption in a number of contemporaneous African countries. Leemann and Mares (2011) and Mares (2015, chapter 6) study bills demanding reform of balloting rules in Prussia and link the demand for secrecy to district characteristics.

smaller in polities which have undergone significant modernization in the years before ballot reform than in polities that have not. This would also be true if there is a switch to negative turnout buying after the reform. If, on the other hand, turnout is encouraged by the secret ballot because of changes in expressive benefits, positive turnout buying or the development of broker-mediated forms of clientelism, then the prediction from the model would be that the this increase would be bigger in places where modernization is more progressed. In all cases, the point we want to stress is that modernization reduces turnout under open voting by reducing vote buying. We test this on a sample of US states from 1870 to 1950 and on a sample of parliamentary constituencies in Great Britain before and after the Ballot Act of 1872.[25]

4.1 The US states

For the sample of US states, we extend Heckelman's (1995) baseline model for the voter turnout rate in gubernatorial elections in the 48 contiguous US states, and estimate the following panel model:

$$(turnoutrate)_{it} = \delta_i + \eta_t + \beta_0 AB_{it} + \beta_1 AB_{it} * M_{it} + \beta_2 M_{it} + Z_{it}\beta_3 + e_{it}, \quad (17)$$

where i represents a state, t represents elections (from 1870 to 1950), AB_{it} is a dummy variable equal to zero before the Australian ballot and equal to one after, M_{it} is a "modernization" variable of interest (*real income per capita*, *average years of schooling* or *urbanization rate*), Z_{it} is a vector of additional control variables, and δ_i and η_t are state and time fixed effects. The turnout interaction hypothesis predicts that $\beta_0 < 0$ and $\beta_1 > 0$. The hypothesis that modernization under open ballot reduces vote buying implies that $\beta_2 < 0$.

Table 3 reports the results. In column one, we reproduce Heckelman's (1995) finding that the Australian ballot has a significant, negative effect on turnout, consistent with the hypothesis that before the Australian ballot turnout is kept high because of vote buying. Columns two to four report on specifications with the interaction between modernization and secret ballot. To facilitate interpretation, we introduce one interaction term at the time. The coefficient on the interaction term is positive and significant in each specification. In column five, we include all three interactions simultaneously. They are less precisely estimated but

[25] The turnout interaction hypothesis cannot be tested on the Western Europe plus off-shoots sample because the turnout data that exist are not comparable across time and space. The main problem is that plural voting was common and that the recorded number of votes does not correspond to the number of voters casting valid vote (the United Kingdom is the leading example of this). Another problem is that open voting by "show of hands" (used e.g., in parts of Denmark till 1901) as opposite to an open ballot system based on color cards (used in many US states) makes it virtually impossible to obtain credible turnout data for the majority of European countries in the sample.

two of them remain significant at the ten percent level and they are jointly significant at the one percent level. The turning points for the modernization variables are all within sample and are indicated at the bottom of the table along with the observed range for each variable. We can, for example, imagine a state which starts with a low education level but expands education over the years. If this state happened to introduce the Australian ballot early, the consequence would be a fairly large drop in turnout, but if it adopts later the drop will be smaller and eventually turnout might even increase. The last effect is not explained by our model because the expressive benefit of voting is constant. In practice this benefit may increase when voting becomes secret. This produces a countervailing effect that could dominate in cases where the importance of the pre-Australian ballot vote market has already been curtailed by modernization. We observe that the point estimate on two of the modernization variables, *real income per worker* and *urbanization rate* are negative. This supports the view that modernization undermined the vote market prior to the Australian ballot. We note, however, that the positive point estimate on *average years of schooling* is inconsistent with this. We conjecture that this variable is picking up a positive correlation between the expressive value of voting and education that dominates any effect running through the vote market. Overall, these results corroborate the hypothesis that modernization undermines the vote market.

Table 3: The effect of the interaction between modernization and secret ballot on the turnout rate in 48 US states 1870-1950.
Dependent variable: *Turnout rate*

	(1)	(2)	(3)	(4)	(5)
Australian ballot	-5.589*	-164.3***	-27.37***	-7.896*	-133.3**
	[-1.96]	[-4.01]	[-3.148]	[-1.85]	[2.29]
Australian ballot* real income per worker		17.50***			12.53*
		[4.13]			[1.76]
Australian ballot* average years of schooling			5.135***		3.07
			[3.23]		[1.59]
Australian ballot* urbanization rate				0.15*	0.005
				[1.68]	[0.04]
Average years of schooling		4.31***	1.84	3.81**	3.246**
		[2.84]	[1.13]	[2.09]	[2.06]
Real income per worker		-20.48***	-9.83**	-13.53***	-15.76**
		[-4.48]	[-2.55]	[-3.50]	[-2.27]
Urbanization rate		-0.49**	-0.54***	-0.50***	-0.56**
		[-2.50]	[-2.77]	[-2.78]	[-2.58]
Women's suffrage	-13.33***	-11.56***	-9.53**	-11.33***	-10.28***
	[-4.23]	[-3.47]	[-2.43]	[-3.00]	[-3.05]
Poll tax	-19.04***	-14.61***	-15.53***	-15.85***	-13.98***
	[-6.03]	[-5.54]	[-5.53]	[-5.19]	[-5.10]
Literacy test	-8.29***	-4.62*	-5.63**	-6.44**	-4.59*
	[-2.87]	[-1.68]	[-2.05]	[-2.16]	[-1.72]
Joint significance of the interaction terms					5.34***
R-squared	0.79	0.82	0.82	0.81	0.83
Turning point	n/a	9.38	5.33	52.64	Not reported

Notes: Robust z-statistics correcting for clustering by state in brackets (i.e., we allow the errors to be correlated over time within states, but not between states). State and year fixed effects included. Number of states is 48; Total number of observations are 1408. Specification (1) corresponds to the baseline model in Heckelman (1995, table 4). *** p<0.01, ** p<0.05, * p<0.1.

4.2 The Ballot Act of 1872

The Second Reform Act of 1867 was followed five years later, in 1872, by the Ballot Act which introduced the Australian ballot in elections to the British House of Commons. We use the Ballot Act as a quasi-natural experiment.[26] We study electoral turnout patterns at the constituency level in the general elections in 1868 and 1874, both of which were conducted according to the new franchise and district rules laid down in 1867, but under different ballot conditions. To test the turnout interaction hypothesis, we proxy modernization with the number of inhabitants per house in each constituency ($density_{it}$) which is closely related to the degree of urbanization.[27] The outcome variable ($turnout_{it}$) records the number of voters who voted in constituency i in election t for $t \in \{1868,1874\}$. We consider the following model for $turnout_{it}$:

$$turnout_{it} = \alpha_i + \nu_t + \alpha_1 R_t + \alpha_2 density_{it} R_t + \alpha_3 density_{it} + \alpha_4 electors_{it} + \varepsilon_{it}, \quad (18)$$

where $R_{1868} = 0$ and $R_{1874} = 1$ indicate that the reform of the ballot took place between 1868 and 1874, α_i is a constituency specific fixed effect, ν_t is an aggregate time effect, $electors_{it}$ is the number of registered voters. We allow $density_{it}$ to have a direct impact on turnout, in addition to its interaction with the ballot reform. Taking the first difference of equation (18), we get

$$\Delta turnout_{it} = (\nu_{1874} - \nu_{1868} + \alpha_1) + \alpha_2 density_{i1874} + \alpha_3 \Delta density_{it} + \alpha_4 \Delta(electors)_{it} + (\varepsilon_{i1874} - \varepsilon_{i1868}). \quad (19)$$

Clearly, the direct effect of the reform (α_1), which we expect to be negative, cannot be identified independently of the common time effect. However, we can recover the interaction effect between the reform dummy variable and the indicator of modernization and test whether $\alpha_2 > 0$. We can also test whether turnout was lower in places that experienced an increase in population density between the two elections, i.e., if $\alpha_3 < 0$. Finally, we expect that the number of electors affects turnout positively, i.e. $\alpha_4 > 0$.

Unfortunately, we only know the votes polled for each candidate and in constituencies

[26] Berlinski and Dewan (2011) use a similar design to study the rise of the Liberal Party in Britain after 1867.

[27] To check the validity of using population density as a measure of urbanization and modernization more generally, we have undertaken three tests. First, in so far as population density and economic development are related, we would expect a positive correlation between population growth in the past and population density. In our sample, this correlation is extremely strong. This demonstrates that the constituencies with the highest population density were those which had grown most in the preceding decade. Second, our sample of one-seat constituencies is concentrated amongst the smaller constituencies. We have checked that the average population density in our sample (5.6) is lower than the density in a sample of towns with more than 50.000 inhabitants (which, excluding London, is 6.0). Third, we have checked using data from the 1871 Census that population density in "rural towns" with less than 5,000 inhabitants is lower than in "urban towns" with more than 50,000 inhabitants.

with more than one seat, voters could cast as many votes as there were seats (Craig, 1977). Consequently, the number of votes is not, in general, equal to the number of voters. However, for the constituencies with only one seat, the number of votes *is* equal to the number of voters who turned out to vote, and our test is based on a sub-sample of 65 one-seat constituencies. The estimation results are

$$\Delta turnout_{it} = \underset{(-1.57)}{-263} + \underset{(1.92)}{61.5} density_{i1874} - \underset{(-2.09)}{614}\,\Delta density_{it} + \underset{(5.31)}{0.47}\Delta electors_{it}, \quad (20)$$

where the figures in parentheses are (robust and clustered) z-statistics. The combined estimate of the reform and the aggregate time effect is negative, consistent with a post-reform drop in turnout, but not statistically significant. More importantly, the coefficient on *density* is positive with a p-value of 0.06. Consistent with the turnout interaction hypothesis, this suggests that the vote market, before the reform, was more vibrant in places that were less densely populated. This conclusion is supported by the negative coefficient on change in *density* which indicates that urbanization undermined the voting market and caused turnout to fall. Overall, the results are consistent with the causal mechanism suggested by our theory and they are corroborated by the insightful case study by Stokes et al. (2013, chapter 8) which concludes that the reason why political parties in Britain adopted legislation to limit electoral corruption and vote buying was that industrialization had made vote buying less economical.[28]

5 Alternative mechanisms and theories

All our empirical tests are consistent with the hypothesis that modernization paves the way for secret ballot. Our theory emphasizes the interaction between vote markets and modernization as the underlying causal mechanism. This is, however, not the only possibility and in this section, we evaluate the plausibility of alternatives channels through which modernization might have triggered ballot reform.

First, while we stress the effect of modernization on demand and supply in a competitive market for votes, Stokes et al. (2013, pp. 179-180) emphasize the importance of brokers, such as party agents, and the way modernization may reduce the value of clientelism. In particular, they explain that "poverty may increase the returns to clientelism by making voters (who on average will have larger marginal utilities of income in poorer

[28] Kam (2013) studies the prices of votes in Britain and finds that they fell after the secret ballot was introduced in 1872. This suggests that vote buying became less important as a consequence of the reform.

societies) more responsive to transfers, whereas growth of average income weakens these returns and thus increases the incentives of brokers to extract rents." Given the data marshalled for this project, it is, however, not possible to discriminate effectively between the vote buying mechanism and broker-mediation mechanism. Both mechanisms are plausible and both predict a positive relationship between secret ballot and modernization and, as such, our evidence is consistent with both and we view them as complementary rather than as substitutes. Modernization may also affect the old regime elites in other ways than proposed by our theory where they are "reluctant" reformers. In Prussia, for example, conservative elements of the elite supported secret ballot in the 19th century presumably because "intimidation of voters in the rural districts of the East was more than matched by the SPD intimidation of shopkeepers, artisans and non-Socialist voters in the cities of the West" (Retallack, 1988, p. 164). In fact, between 1869 and 1912, 18 bills were proposed in the Reichstag to introduce elements of secrecy in voting. Mares (2015, chapter 6) shows that bills were sponsored by the politicians who were elected in districts with a high degree of occupational fragmention. This suggests that economic development through its effect on economic diversity created demand from within the German elite for ballot reform. Along similar lines Elklit (1988, p. 300) notes that voters of all parties, including the conservative and social democratic party, experienced electoral coercion in Denmark prior to the secret ballot. Depending on shifts in the relative ability of parties to coerce voters, subsets of the elite might, therefore, see ballot reform as an expedient way to gain electoral advantage. This line of reasoning suggests an alternative way that modernization may drive ballot reform not present in our model: modernization empowers the working class and create economic diversification which make the old elite favor ballot reform simply because intimidation becomes unworkable. One way to evaluate the power of this mechanism in our context is to include the share of left-wing party seats in our cross-country model for secret ballot adoption and test whether it has a positive effect. We find that it does, but the effect is statistically insignificant. Since the coefficient on GDP per capita remains positive and significant this demonstrates that modernization must work through other channels than expedient elite support for reform.[29]

Second, another consequence of modernization, which is stressed by, for example, Kitschelt and Wilkingson (2007) and Keefer and Vlaicu (2008), is the emergence of "modern" political parties. That is, parties with either credible policy platforms which

[29] The social democratic parties in both Denmark and Germany were in favor of secret ballot, see Elklit (1988, p. 302) and Schorske (1955, p. 3). The results are reported in the supplementary appendix Table S5.

individual politicians respect or with organizational structures that enable monitoring of new forms of exchange (other than outright vote buying) between the party and the voters. Once a modern party is established, the value of open voting is reduced, either because the party can contest elections with programmatic policy platforms or because it can support forms of clientelism that do not require open voting. It is, therefore, possible that it is the emergence of modern political parties that triggers secret ballot, and that modernization operates through this channel. To investigate this, we determine the approximate year in which political parties with nation-wide organizations started to contest elections emerged in our Western Europe plus off-shoots sample. The recorded timing pattern relative to the timing of secret ballot reform provides mixed support for this hypothesis (see supplementary material appendix Table S1). In some countries, modern parties did indeed emerge prior to the introduction of the secret ballot (e.g., in Germany and Denmark), but in others they did not (e.g., in the United Kingdom and the Netherlands). We add a dummy variable capturing the year in which modern political parties emerged to the baseline specification. We find that this variable denoted by ”effective parties” itself has a positive, but insignificant effect on the probability of secret ballot reform. Importantly, the positive effect of the modernization variables remains statistically significant.[30] This suggests that the modernization effect did not primarily operate through the emergence of modern political parties.

Third, the historical narrative from the USA suggests that the old regime elites played a proactive role in ballot reform. In fact, it has been suggested that they used the Australian ballot to disenfranchise illiterates among foreign born citizens and blacks in order to obtain a partisan advantage (Keyssar, 2000, Heckelman, 2000). If this mechanism is important, we would expect higher support for the Australian ballot in states with lower levels of education which would imply higher levels of illiteracy. This does not seem to be borne out by the data as years of schooling is positively correlated with the probability of ballot reform in the 48 US states.

Forth, modernization may encourage the formation of “public opinion” and induce a fall in the cost of mass communication. Stokes et al. (2013) see the later as a consequence of rising literacy and increasing years of schooling. In our analysis of the 48 US states, we find that *GDP per capita* and *average years of schooling* are both significant. This, on the one hand, is consistent with the possibility that modernization works through a fall in the cost of mass communication. On the other hand, this cannot be the whole story as income growth has

[30]The result is reported in the supplementary appendix Table S5.

an independent effect. Moreover, Stokes et al. (2013, p. 238) suggest that campaigning in newspapers did not become common till the first decades of the 20th century. We have, therefore, checked whether the adoptions of secret ballot after 1900 drive the correlation between ballot reform and income in our samples and find that it does not.

Finally, our theory puts a strong emphasis on vote buying and less emphasis on intimidation and coercion. Yet, there are plenty of historical examples of how intimidation and coercion corrupted elections. Our theory offers some scope for discriminating between the two forms of electoral corruption. While the effect on turnout of secret ballot is unambiguously negative if vote markets are important, the effect is less clear if electoral corruption takes the form of intimidation and coercion. In that case, voter turnout may actually increase after the Australian ballot was introduced. In contrast to Heckelman (1995) who finds a clear fall in turnout amongst US states after the vote became secret, the cross-country analysis by Przeworski (2008) shows that the effect can be positive. This suggests that context matters and that both vote buying and intimidation played a role. Importantly, however, whether the direct effect on turnout is negative or positive does not matter for the hypothesis that modernization causes secret ballot adoption. The reason, as argued by Stokes et al. (2013), is that coercion and intimidation are less effective means of controlling the electorate with higher levels of economic development. Accordingly, whether modernization works through the vote market or through a reduction in the returns to coercion and clientelism, the outcome is that modernization makes secret ballot more likely. Our evidence strongly suggest that this was the case.

6 Conclusion

We " unbundle" the concept of democracy in order to evaluate in a more nuanced way the interplay between modernization and democratization. Our event history studies of the introduction of the secret ballot demonstrate a remarkably robust relationship between modernization and its adoption. This finding is important because it grants the forces of urbanization, rising education standards, and income growth a role in explaining political development. The role is more limited than envisaged by Lipset (1959), but we contend that modernization, while probably not causally linked to the timing of the major suffrage reforms and thus to the evolution of the overall package of democratic institutions, was a causal factor in getting the secret ballot introduced.

Uncovering the precise mechanism behind this aggregate result requires more refined work than what we have presented here. We have proposed one mechanism based on the logic of a competitive vote market. The reported evidence on the interplay between modernization, ballot reform and turnout is consistent with this mechanism, but it does not rule out that other forms of clientelism could also have played a major role. More work is clearly needed to gain a fuller understanding of what is behind the relationship between modernization and the introduction of secret ballot.

References

Acemoglu, Daron., Robinson, James A., 2006. Economic origins of dictatorship and democracy. Cambridge University Press, Cambridge, UK.

Acemoglu, Daron., Johnson, Simon., Robinson, James A., Yared, Pierre., 2008. Income and democracy. American Economic Review 98(3), 808-842.

Aidt, Toke S., Jensen, Peter S. 2009. The taxman tools up: An event history study of the introduction of the personal income tax, 1815-1938. Journal of Public Economics 93, 160-175.

Aidt, Toke S., Giovannoni, Francesco. 2011. Constitutional rules. Social Choice and Welfare 37, 219-268.

Aidt, Toke S., Jensen, Peter S. 2014. Workers of the world, unite! Franchise extensions and the threat of revolution in Europe, 1820-1938. European Economic Review 72, 52-75.

Aidt, Toke S., Franck, Raphael. 2015. Democratization under the threat of revolution: Evidence from the Great Reform Act of 1832. Econometrica 83(2),505-547..

Aidt, Toke S., Leon, Gabriel. 2015. The democratic window of opportunity: Evidence from riots in sub-Saharan Africa. Journal of Conflict Resolution, forthcoming.

Alvarez, Michael E., Cheibub, Jose A., Limongi, Fernando, Przeworski, Adam (1996). Classifying political regimes. Studies in Comparative International Development, 31(2), 3-36.

Anderson, Gary M., Tollison, Robert D., 1990. Democracy in the marketplace. In William M. Crain and Robert D. Tollison (eds.), Predicting Politics: Essays in Empirical Public Choice, 285-303. Ann Arbor: University of Michigan Press.

Andrews, Josephine T., Jackman, Robert W., 2005. Strategic fools: Electoral rule choice under extreme uncertainty. Electoral Studies 24(1), 65-84.

Angrist, Joshua, Pischke, Jörn-Steffen, 2009. Mostly harmless econometrics. Princeton University Press.

Ansell, Ben, Samuels, David, 2010. Inequality and democratization: A contractarian approach. Journal of Comparative Politics 20(10), 1-32.

Ansell, Ben, Samuels, David, 2015. Inequality and Democratization: An Elite-Competition Approach. Cambridge University Press, U.K.

Baland, Jean-Marie., Robinson, James.A., 2007. How does vote buying shape the economy. In Frederic C. Schaffer (ed.). Elections for Sale: The Causes and Consequences of Vote Buying, 123-141. Lynne Rienner Publishers, Boulder, Colorado.

Baland, Jean-Marie., Robinson, James.A., 2008. Land and power: Theory and evidence from Chile. American Economic Review 98(5), 1737-1765.

Barro, Robert J., 1999. Determinants of democracy. Journal of Political Economy 107, 6–29.

Beck, Nathaniel, Katz, Jonathan N., Tucker, Richard, 1998. Taking time seriously: Time-series-cross-section analysis with a binary dependent variable. American Journal of Political Science 42(4), 1260–1288.

Berlinski, Samuel., Dewan, Torun., 2011. The Political Consequences of Franchise Extension: Evidence from the Second Reform Act. Quarterly Journal of Political Science 6(34), 329-376.

Bensel, Richard, 2004. American Ballot Box in the Mid-Nineteenth Century: Law, Identity and the Polling Place. Cambridge University Press, Cambridge, UK.

Blais, André, Dobrzynska, Agnieska, Indridason, Indridi H., 2004. To adopt or not to adopt proportional representation: The politics of institutional choice. British Journal of Political Science 35, 182–190.

Boix, Carles, 1999. Setting the rules of the game: The choice of electoral systems in advanced democracies. American Political Science Review 93, 609-624.

Boix, Carles, 2003. Democracy and redistribution. Cambridge University Press, Cambridge, UK.

Boix, Charles, 2011. Democracy, development, and the international system. American Political Science Review, 105 (4), 809-828.

Braun, Sebastian, Kvasnicka, Michael, 2013. Men, women, and the ballot. Gender imbalances and suffrage extensions in US states. Explorations in Economic History, 50 (3), 405–426.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. Bootstrap-Based Improvements for Inference with Clustered Errors. Review of Economics and Statistics, 90(3): 414--27.

Congleton, Roger.D., 2011. Perfecting parliament. Cambridge University Press, Cambridge, UK.

Converse, Philip E. 1974. Comment on Burnham's "Theory and Voting Research." American Political Science Review 68, 1024-1027.

Collier, Paul, Vicente, P.C., 2012. Violence, bribery, and fraud: the political economy of elections in Sub-Saharan Africa. Public Choice, 153, 117-147.

Cox, Gary W. and Kousser, J. Morgan, 1981. Turnout and Rural Corruption: New York as a Test. American Journal of Political Science 25(4) Nov., 646-663.

Craig, Fred W.S., 1977. British parliamentary election results. Macmillan, Great Britain.

Dahl, Robert, 1971. Polyarchy: Participation and opposition. Yale University Press, New Haven.

Dal Bo, Ernesto, 2007. Bribing Voters. American Journal of Political Science 51(4), 789-803.

Dekel, Eddie, Jackson, Matthew O., Wolinsky, Asher, 2008. Vote Buying: General Elections. Journal of Political Economy 116(2), 351-380.

Elklit, Jørgen, 2000. Free and fair elections. In R. Rose (ed.). International encyclopedia of elections, 130-134, CQ Press, Washington, D.C.

Fearon, James, 2003. Ethnic and cultural diversity by country. Journal of Economic Growth 8, 195-222.

Fredman, Lionel E., 1968. The Australian ballot. The story of an American reform. Michigan University Press, Ann Arbor, Michigan.

Galor, Oded, Moav, Omer, Vollrath, Dietrich, 2009. Inequality in landownership, the emergence of human-capital promoting institutions, and the great divergence. Review of Economic Studies 76(1), 143–179.

Gleditsch, Kristian S., Ward, Michael D., 2006. Diffusion and the international context of democratization. International Organization 60(4), 911-933.

Gundlach, Erich, Paldam, Martin, 2009. A farewell to critical junctures, European Journal of Political Economy 25(3), 340-354.

Kam, Christopher, 2013. The impact of the secret ballot on bribe prices at nineteenth-century parliamentary elections. Unpublished paper.

Heckelman, Jac C., 1995. The effect of the secret ballot on voter turnout rates. Public Choice 82(1/2), 107-124.

Heckelman, Jac C., 2000. Revisiting the Relationship Between Secret Ballots and Turnout: A New Test of Two Legal-institutional Theories. American Politics 28, 194-215.

Heckelman, Jac C, Yates, Andrew J., 2002. Incumbency preservation through electoral legislation: The case of the secret ballot. Economics of Governance 3(1), 47-57

Keefer, Philip, Vlaicu, Razvan (2008). Democracy, Credibility, and Clientelism Journal of Law, Economics and Organization 24 (2): 371-406.

Hicken, Allen D., 2007. How do rules and institutions encourage vote buying? In Frederic C. Schaffer (ed.). Elections for Sale: The Causes and Consequences of Vote Buying, 33-60. Lynne Rienner Publishers, Boulder, Colorado.

Keyssar, Alexander, 2009. The Right to Vote. The Contested History of Democracy in the United States. Basic Books, New York.

King, Gary, Zeng, Langche, 2001. Logistic regression in rare events data. Political Analysis 9, 137–163.

Kitschelt, H., Wilkinson, W.I. (2007), Patrons, Clients and Policies: Patterns of Democratic Accountability and Political Competition. Cambridge University Press, Cambridge.

Kuo, Didi, Teorell, Jan, 2013. Election Fraud, Ballot Reform and Contested Congressional Elections: An Analysis of the United States, 1860-1930. Unpublished working paper.

Lee, Stephen J. 1994. Aspects of British Political History 1815-1914. Routledge, London and New York.

Leemann, Lucas, Mares, Isabella, 2011. From 'open secrets' to the secret ballot: The economic and political determinants of secret ballot reform. Unpublished working paper. University of Columbia.

Leeson, Peter, Dean, Andrea, 2009. The Democratic Domino Theory: An Empirical Investigation. American Journal of Political Science 53(3), 533-551.

Lehoucq, Fabrice, 2007. When does a market for votes emerge. In Frederic C. Schaffer (ed.). Elections for Sale: The Causes and Consequences of Vote Buying, 33-60. Lynne Rienner Publishers, Boulder, Colorado.

Lipset, Seymour M., 1959. Some social requisites of democracy: Economic development and political legitimacy. American Political Science Review 53(1), 69-105.

Lott, John R., Kenny, Lawrence W., 1999. Did women's suffrage change the size and scope of government? Journal of Political Economy 107(6), 1163-1198.

Ludington, Arthur C. American ballot laws, 1888-1910, Bibliolife, United Kingdom.

Mackie, Thomas T., 2000. The secret ballot. In Rose, Richard (ed.). International encyclopedia of elections, , 289-290. CQ Press, Washington, D.C.

Mares, Isabela, 2015. From Open Secrets to Secret Voting: Democratic Electoral Reforms and Voter Autonomy. Cambridge University Press, New York, New York.

McCook, J.J., 1892. The alarming proportion of venal voters, Forum XIV, 1-13.

Miller, Michael K. 2012. "Economic Development, Violent Leader Removal, and Democratization." American Journal of Political Science 56(4), 1002–20.

Moore, Barrington, 1966. The social origins of democracy and dictatorship. Beacon, Boston, MA.

Ostrogorski, Moisei, 1964. Democracy and the organization of political parties. Quadrangle Books, Chicago, IL.

Persson, Torsten, Tabellini, Guido, 2009. Democratic capital: the nexus of political capital and economic Change. American Economic Journal 1, 88-126.

Przeworski, Adam, 2008. Constraints and choices. Electoral participation in historical

perspective. Comparative Political Science 42(1), 4-30.
Przeworski, Adam, 2009. Conquered or granted? A history of suffrage extensions. British Journal of Political Science 39, 291-321.
Przeworski, Adam, 2010. Suffrage and Voting Secrecy in General Elections. Unpublished Working Paper.
Przeworski, Adam, Limongi, Fernando., 1997. Modernization: Theories and facts. World Politics 49(2), 155-183.
Przeworski, Adam, Alvarez, Michael E., Cheibub, Jose A., Limongi, Fernando, 2000. Democracy and development: Political institutions and well-being in the world, 1950-1990. Cambridge University Press, Cambridge, UK.
Retallack, James, 1988. Notables of the right. The Conservative Party and political mobilization in Germany, 1876-1918. Unwin Hyman, Boston, MA.
Robinson, James and Verdier, T. 2013. The political economy of clientelism. Scandinavian Journal of Economics 115(2), 260–291.
Rokkan, Stein, 1961. Mass suffrage, secret voting and political participation. European Journal of Sociology 2, 132-152.
Rueschemeyer, Dietrich, Stephens, Evelyne H., Stephens, John D., 1993. The impact of economic development on democracy. The Journal of Economic Perspectives 7(3), 71-86.
Rusk, Jerrold G. 1974. Comment: The American electoral universe: Speculation and evidence. American Political Science Review 68, 1028-1049.
Schofield, Philip, 2004. Jeremy Bentham, the French Revolution and political radicalism. History of European Ideas 30, 381-401
Schorske, Carl, 1955. German Social Democracy, 1905-1917. Harvard University Press, Cambridge, MA.
Seymour, Charles, 1915. Electoral reform in England and Wales. Yale University Press, New Haven, CT.
Seymour, Charles., Frary, Donald P., 1918. How the world votes. C.A. Nichols Company.
Snyder, James M., 1991. On Buying Legislatures. Economics and Politics 3(2), 93-109.
Stokes, Susan, 2005. Perverse accountability; a formal model of machine politics with evidence from Argentina. American Political Science Review 99(3), 315-325.
Stokes, Susan, Dunning, Thad, Nazareno, Marcelo, Brusco, Valeria, 2013. Brokers, Voters, and Clientelism: puzzle of distributive politics. Cambridge University Press, Cambridge, UK.
Tilly, Charles, 1993. European revolutions: 1492-1992. Wiley-Blackwell, Oxford, UK.
Treisman, Daniel, 2015. Income, Democracy, and Leader Turnover. American Journal of Political Science, 59 (4), 927-942.
Wejnert, Barbara, 2005. Diffusion, development, and democracy, 1800-1999. American Sociological Review 70, 53-81.
Ziblatt, Daniel, 2008. Does landholding inequality block democratization? A test of the " bread and democracy" thesis and the case of Prussia. World Politics 60(4), 610-641.
Ziblatt, Daniel, 2009. Shaping democratic practice and the causes of electoral fraud: The case of nineteenth-century Germany. American Political Science Review 103(1), 1-21.

Supplementary material

(For online publication)

Appendix S1: Proofs and discussion of assumptions

One party only buys votes. The two first order conditions are

$$n^b_{j,\sim j}: \frac{\partial g}{\partial v}(v) \leq \frac{v_j \hat{\eta} \Delta}{\lambda_{\sim j}\alpha_j(\Delta + M)} + \frac{\Delta}{2\alpha_j} \equiv K_j. \tag{21}$$

The left-hand side is the same. If $K_E \neq K_R$, then at most one of these equations can be satisfied with equality. Since the function g is strictly concave (so that $\frac{\partial g}{\partial v}(v)$ is decreasing in v), it is the equation associated with the minimum of K_E and K_R that may admit an interior solution. Under Assumption 1, $K_E < K_R$ and $n^{b^*}_{ER} > 0$ and $n^{b^*}_{RE} = 0$. The second order conditions, including

$$\Upsilon \equiv \frac{\Delta + M}{\hat{\eta}} \alpha_E^2 \frac{\partial^2 g}{\partial v^2} < 0 \tag{22}$$

are satisfied.

The continuation values of veto and reform. To calculate the continuation payoff for party E following a veto, let $V(OB, E)$ be the expected present value for party E at the beginning of a period if the present ballot regime is OB and it wins the next election, and let the corresponding expected present value if it loses be $V(OB, R)$. Under the assumption that party E will veto when needed, we can write the value of calling a veto at the end of the current period as follows:

$$W_E(\text{veto}) = -\rho + \beta f_E^{OB} V(OB, E) + \beta(1 - f_E^{OB}) V(OB, R), \tag{23}$$

where

$$\begin{aligned} V(OB, E) &= u_E(E) + M - v_E \Delta \lambda_R^{-1} n^{b^*}_{ER}(.) + \beta f_E^{OB} V(OB, E) \\ &\quad + \beta(1 - f_E^{OB}) V(OB, R), \\ V(OB, R) &= u_E(R) - v_E \Delta \lambda_R^{-1} n^{b^*}_{ER}(.) - \rho + \beta f_E^{OB} V(OB, E) \\ &\quad + \beta(1 - f_E^{OB}) V(OB, R). \end{aligned}$$

We can solve these two equations for $V(OB, E)$ and $V(OB, R)$, which yields

$$V(OB, E) = \frac{(1 - \beta(1 - f_E^{OB})(u_E(E) + M) + \beta(1 - f_E^{OB})(u_E(R) - \rho) - v_E \Delta \lambda_R^{-1} n^{b^*}_{ER}(.)}{1 - \beta},$$

$$V(OB, R) = \frac{(1 - \beta f_E^{OB})(u_E(R) - \rho) + \beta f_E^{OB}(u_E(M) + M) - v_E \Delta \lambda_R^{-1} n^{b^*}_{ER}(.)}{1 - \beta}.$$

Substituting these two equations into equation (23) and simplifying give equation (14) in the text. The continuation value following a permanent transition to secret ballot can be constructed in a similar fashion, but we may simply note that the per period expected utility of the elite is equal to

$$(1 - f_E^{SB}) u_E(R) + f_E^{SB}(u_E(E) + M)$$

and that this is repeated for all periods starting one period ahead and is, therefore, discounted back to the present by $\frac{\beta}{1-\beta}$. This gives equation (13) in the text.

Comparative statics. Total differentiation of equation (10) yields:

$$\frac{\partial n_{ER}^{b^*}}{\partial \alpha_E} = -\frac{\frac{1}{\hat{\eta}}\frac{\partial g}{\partial v}(\Delta + M)}{\Upsilon} > 0; \frac{\partial n_{ER}^{b^*}}{\partial \lambda_R} = -\frac{\Delta v_E}{\lambda_R^2 \Upsilon} > 0; \frac{\partial n_{ER}^{b^*}}{\partial N_R} = 0. \tag{24}$$

Substituting $n_{ER}^{b^*}$ along with the equilibrium values of electoral turnout from equation (6) into equation (12) gives

$$f_E^{OB} = f_E(n_{ER}^{b^*}(\alpha_E, \lambda_R), 0, n_E^*(0, N_E), n_R^*(n_{ER}^{b^*}(\alpha_E, \lambda_R), N_R)). \tag{25}$$

The comparative statics results are

$$\frac{\partial f_E^{OB}}{\partial \alpha_E} = \frac{1}{\hat{\eta}}\left[\frac{\Delta v_E \hat{\eta}}{\lambda_R(\Delta + M)}\right]\frac{\partial n_{ER}^{b^*}}{\partial \alpha_E} + \frac{1}{\hat{\eta}}\frac{\partial g}{\partial v} n_{ER}^{b^*} > 0, \tag{26}$$

$$\frac{\partial f_E^{OB}}{\partial \lambda_R} = \frac{1}{\hat{\eta}}\left[\frac{\Delta v_E \hat{\eta}}{\lambda_R(\Delta + M)}\right]\frac{\partial n_{ER}^{b^*}}{\partial \lambda_R} > 0, \tag{27}$$

$$\frac{\partial f_E^{OB}}{\partial N_R} = \frac{1}{\hat{\eta}}\left[\frac{-1}{2}\frac{\partial n_R^*}{\partial n_{ER}^{b^*}} + \alpha_E \frac{\partial g}{\partial v}\right]\frac{\partial n_{ER}^{b^*}}{\partial N_R} - \frac{1}{2\hat{\eta}}\frac{\partial n_R^*}{\partial N_R} = -\frac{1-c}{2\hat{\eta}} < 0, \tag{28}$$

where we, in each case, use the first order condition for $n_{ER}^{b^*}$ and the relevant comparative statics from equation (24) to sign the effects. Rewrite equation (15) and define

$$\begin{aligned} D(\alpha_E, \lambda_R, N_R) \equiv \beta(f_E^{OB} - f_E^{SB})(\Delta + M) - \beta v_E \Delta \lambda_R^{-1} n_{ER}^{b^*}(.) \\ - (1 - \beta f_E^{OB})\rho. \end{aligned} \tag{29}$$

Since the elite vetoes only if $D(\alpha_E, \lambda_R, N_R) \geq 0$, any factor that reduces $D(.)$ makes a veto less likely and ballot reform more likely. The comparative statics are

$$\begin{aligned} \frac{\partial D}{\partial \alpha_E} &= \beta \frac{\partial f_E^{OB}}{\partial \alpha_E}(\Delta + M) - \beta v_E \Delta \lambda_R^{-1} \frac{\partial n_{ER}^{b^*}}{\partial \alpha_E} + \beta \frac{\partial f_E^{OB}}{\partial \alpha_E}\rho \\ &= \beta(\Delta + M)\frac{1}{\hat{\eta}}\frac{\partial g}{\partial v} n_{ER}^{b^*} + \beta \frac{\partial f_E^{OB}}{\partial \alpha_E}\rho > 0 \end{aligned} \tag{30}$$

$$\begin{aligned} \frac{\partial D}{\partial \lambda_R} &= \beta \frac{\partial f_E^{OB}}{\partial \lambda_R}(\Delta + M) - \beta v_E \Delta \lambda_R^{-1} \frac{\partial n_{ER}^{b^*}}{\partial \lambda_R} + \beta v_E \Delta n_{ER}^{b^*} \lambda_R^{-2} + \beta \frac{\partial f_E^{OB}}{\partial \lambda_R}\rho \\ &= \beta v_E \Delta n_{ER}^{b^*} \lambda_R^{-2} + \beta \frac{\partial f_E^{OB}}{\partial \lambda_R}\rho > 0, \end{aligned} \tag{31}$$

where we use the first order condition for $n_{ER}^{b^*}$ and the results in equations (26) to (27) to get the second line of each expression. Moreover, we find that

$$\frac{\partial D}{\partial N_R} = \beta(\Delta + M)\left(\frac{\partial f_E^{OB}}{\partial N_R} - \frac{\partial f_E^{SB}}{\partial N_R}\right) + \beta \frac{\partial f_E^{OB}}{\partial N_R}\rho = \beta \frac{\partial f_E^{OB}}{\partial N_R}\rho < 0 \tag{32}$$

since $\frac{\partial f_E^{SB}}{\partial N_R} = \frac{\partial f_E^{OB}}{\partial N_R} = -\frac{1-c}{2\hat{\eta}}$ because $\frac{\partial n_{ER}^{b^*}}{\partial N_R} = 0$.

Simplifying assumptions.

1. Voter type is observable. We assume that the parties can observe who the opposition voters are and so offer them their reservation price in exchange for their vote. The other extreme is to assume that the parties cannot observe who is who but know the reservation prices of the various types and so that $p_R < p_E$. In stage C, party E can offer p_R. At this price only voters of type R will sell. Party E can get them to self-select and so the probability that it offers a bribe to an opposition voter is 1. Party R needs to offer p_E to induce voters of type E

to sell. At this price, both types are willing to sell and so the party R must buy votes at random. The probability that it buys a voter of type E is $\frac{1}{n_R+n_E}$. In stage B, we maintain the assumption that voters expect that everyone shows up when they calculate their own chances of being offered a bribe. A voter of type R expects to get a bribe from party R with probability $\frac{n_{ER}^b}{N_R}$ or from party E with probability $\frac{n_{RE}^b}{N_R+N_E}$ (this assumes he cannot get bribed twice) and his expected bribe income is $p_R\frac{n_{ER}^b}{N_R}+p_E\frac{n_{RE}^b}{N_R+N_E}$. A voter of type E expects to get a bribe from party R with probability $\frac{n_{RE}^b}{N_R+N_E}$. The turnouts of voters of the two groups are:

$$n_E^{OB}=N_E(1-c)+N_E\frac{n_{RE}^b}{N_R+N_E}p_E\lambda_E,$$
$$n_R^{OB}=N_R(1-c)+N_R\lambda_R\left(p_R\frac{n_{ER}^b}{N_R}+p_E\frac{n_{RE}^b}{N_R+N_E}\right),$$

where $n_E^{OB}+n_R^{OB}=(N_R+N_E)(1-c)+\Delta(n_{ER}^b+n_{RE}^b)$.

In stage A, the two parties now internalize both the effect on turnout and on their chance of buying an influential vote. We let

$$v=\alpha_E n_{ER}^b-\alpha_R\frac{n_{RE}^b}{n_R^{OB}+n_E^{OB}}.$$

The first order condition will now have to take into account that v is influenced by total turnout which is influenced by voting buying with

$$\frac{\partial v}{\partial n_{ER}^b}=\alpha_E+\alpha_R\frac{n_{RE}^b}{(n_R^{OB}+n_E^{OB})^2}\Delta,$$
$$\frac{\partial v}{\partial n_{RE}^b}=-\alpha_R\frac{(n_R^{OB}+n_E^{OB})-n_{RE}^b\Delta}{(n_R^{OB}+n_E^{OB})^2}.$$

We write the first order conditions as

$$\frac{\partial v}{\partial n_{ER}^b}\frac{\partial g}{\partial v}\le\alpha_E K_E,$$
$$-\frac{\partial v}{\partial n_{RE}^b}\frac{\partial g}{\partial v}\le\alpha_R K_R,$$

with $K_E<K_R$ defined in Appendix A. Conjecture that only party E will buy. Then $\frac{\partial v}{\partial n_{ER}^b}=\alpha_E$ and $\frac{\partial g}{\partial v}=K_E$. Look at

$$\alpha_R\frac{(n_R^{OB}+n_E^{OB})-n_{RE}^b\Delta}{(n_R^{OB}+n_E^{OB})^2}K_E\le\alpha_R K_R.$$

If $\frac{(n_R^{OB}+n_E^{OB})-n_{RE}^b\Delta}{(n_R^{OB}+n_E^{OB})^2}<1$, which it is if $(n_R^{OB}+n_E^{OB})>1$, then the conjecture is correct and only party E bribes. In this case, all the relevant comparative statics derive from $\frac{\partial g}{\partial v}=K_E$ and are exactly as before.

2. Rational expectations about turnout. If voters believe that their chance of getting a bribe is proportional to turnout rather than to the size of their group, then the expected utility value of the bribe is

$$u_j^e=\frac{n_{\sim jj}^b}{n_j}p_j\lambda_j. \quad (33)$$

At equilibrium the expected turnout must be equal to actual turnout, so turnout for group j is

defined by

$$n_j = N_j\left(1 - c + \frac{n^b_{\sim jj}}{n_j}p_j\lambda_j\right), \tag{34}$$

with the solution[31]

$$n^*_j(n^b_{\sim jj}, N_j) = \frac{1}{2}N_j(1-c) + \frac{1}{2}\sqrt{N_j^2(1-c)^2 + 4N_j n^b_{\sim jj}p_j\lambda_j}. \tag{35}$$

The main difference between this and the case considered in the text is that the marginal effect of bribery on turnout is not independent of $n^b_{\sim jj}$. Under the assumption that

$$\frac{p_R v_E \hat{\eta}}{\alpha_E(\Delta_E + M)} + \frac{1}{2}\frac{\partial n^*_R}{\partial n^b_{ER}} < \frac{p_E v_R \hat{\eta}}{\alpha_R(\Delta_R + M)} + \frac{1}{2}\frac{\partial n^*_E}{\partial n^b_{RE}},$$

it remains true that only party E is active in the vote market. All comparative static results continue to hold as long as vote buying is optimal. The only result that requires an additional assumption is the one regarding the effect of N_R on the net value of a veto.

3. $\Delta_E = \Delta_R$. Since these represent per capita benefits, the natural alternative is that $\Delta_E > \Delta_R$, i.e., each member of the (small) elite stands to lose a lot more than each member of the (large) majority. All results go through when $\Delta_E > \Delta_R$ because this simply increases the reservation price for elite voters. If, for some reason, $\Delta_E < \Delta_R$, all results holds as long as this difference is not sufficient to violate Assumption 1.

4. Secret ballot as an absorbing state. Acemoglu and Robinson (2000) make a similar assumption about franchise extensions to capture the notion that institutional reforms persist. If we allow for reversals, it is possible that there exist equilibria where the society fluctuates between the two ballot systems (this would for example be the case if $\rho = 0$), but there will also exist parameter values such that a transition to secret ballot is permanent. It is a question of finding parameters such that party E will not veto, while party R will.

5. The vote buying technology. This is the critical assumption that leads to the asymmetric outcome in the vote market with at most one party buying. Other technologies lead to situations where both parties buy votes. This makes the analysis more blurred, but as long as one of the parties has a clear advantage under open ballot and the other has an advantage under secret ballot, the logic of reform continues to apply and modernization continues to undermine the value of the vote market by pushing up prices and increasing transaction costs.

6. Buying back votes. In the main specification, we allow parties to target opposition voters, but not to 'buy back' supporters who have been offered a vote contract by the opposition. The logic of the model is not substantially affected by this restriction. The maintained assumption in the model is that vote contracts can be enforced, i.e., that the party that buys a vote can penalize the seller if the contract is broken. This means that a party that buys a vote back will have to compensate the voter for this cost. If we let the utility cost of the penalty that party j can impose on contracts which are broken be q_j. We need to introduce some extra notation. Let n^b_{jj} be the number of voters of type j that party j buys back with $n^b_{jj} \leq n^b_{\sim jj}$ and let the bribe prices that party R offers to voters of type E be $p^R_E = \lambda_E^{-1}\Delta$, the bribe price that R offers to own voters to buy them back $p^R_R = \lambda_R^{-1}q_E$ and likewise for party E. The expected turnout of voters of type j is

$$n^{OB}_j = N_j(1-c) + \left(n^b_{\sim jj} - n^b_{jj}\right)p_j^{-j}\lambda_j + n^b_{jj}p_j^j\lambda_j \tag{36}$$

[31] We disregard the negative root of the polynomial.

$$= N_j(1-c) + \left(n^b_{\sim jj} - n^b_{jj}\right)\Delta + n^b_{jj} q_{\sim j}$$

and the vote productivity adjusted vote shares

$$v = \alpha_E(\max\left[n^b_{ER} - n^b_{RR}, 0\right]) - \alpha_R(\max\left[n^b_{RE} - n^b_{EE}, 0\right]). \quad (37)$$

The objective functions of the two parties are

$$\begin{aligned} &f_E(\Delta + M) - p^E_R v_E \max\left[n^b_{ER} - n^b_{RR}, 0\right]) - p^E_E v_E n^b_{EE} \\ &-f_E(\Delta + M) - p^R_E v_R \max\left[n^b_{RE} - n^b_{EE}, 0\right]) - p^R_R v_R n^b_{RR}. \end{aligned} \quad (38)$$

The first order condition associated with the four choice variables can be written as

$$n^b_{ER}: \frac{1}{\hat{\eta}}\frac{\partial g}{\partial v}(v) \leq \frac{v_E \Delta}{\lambda_R \alpha_E(\Delta + M)} + \frac{\Delta}{2\hat{\eta}\alpha_E} \text{ with } = \text{if } n^b_{ER} > 0 \quad (39)$$

$$n^b_{EE}: \frac{1}{\hat{\eta}}\frac{\partial g}{\partial v}(v) \leq \frac{v_E q_R}{\lambda_E \alpha_R(\Delta + M)} + \frac{(q_R - \Delta)}{2\hat{\eta}\alpha_R} \text{ with } = \text{if } n^b_{EE} > 0 \quad (40)$$

$$n^b_{RE}: \frac{1}{\hat{\eta}}\frac{\partial g}{\partial v}(v) \leq \frac{v_R \Delta}{\lambda_E \alpha_R(\Delta + M)} + \frac{\Delta}{2\hat{\eta}\alpha_R} \text{ with } = \text{if } n^b_{RE} > 0 \quad (41)$$

$$n^b_{RR}: \frac{1}{\hat{\eta}}\frac{\partial g}{\partial v}(v) \leq \frac{v_R q_E}{\lambda_E \alpha_E(\Delta + M)} + \frac{(q_E - \Delta)}{2\hat{\eta}\alpha_E} \text{ with } = \text{if } n^b_{RR} > 0 \quad (42)$$

Under assumption A1, $n^b_{ER} > 0$ and $n^b_{RE} = 0$. That is, party E bribes voters of type R, but party R does not bribe voters of type E. Clearly, $n^b_{EE} = 0$. However, it is possible that $n^b_{RR} > 0$, i.e., that party R will bribe some of its voters, who have been targeted by party E, to get them to shift their vote back to it. This is ruled out if

$$\frac{v_E \Delta}{\lambda_R(\Delta + M)} + \frac{\Delta}{2\hat{\eta}} < \frac{v_R q_E}{\lambda_E(\Delta + M)} + \frac{(q_E - \Delta)}{2\hat{\eta}}. \quad (43)$$

This condition is satisfied if the sanction that party E can impose for breaking the initial vote contract between it and a voter of type R (q_E) is sufficiently large, and/or if the cost of raising funds for party E is sufficiently low relative to party R or if the marginal value of income is much lower for voters of type E than for voters of type R. In this case, all the results stated in the text hold true. If the condition fails, the vote market will shot down, as party E will stop buying votes because it foresees that party R will buy them back. We notice that modernization may induce this by lowering q_E – the sanction that party E can impose for a broken vote contract may become less costly as alternative employment opportunities opens up for voters of type R.

Appendix S2: Timing of the secret ballot

For the timing of the secret ballot, we use the year in which the electoral law which introduced de facto secret ballot was passed by parliament.

The Netherlands: 1849. Stuurman (1991, pp. 462-463) notes that " in the autumn of 1849 Thorbecke, the architect of the new constitution, was at last called upon to form a Cabinet and it was his government that produced the final Electoral law [...] There was a secret ballot in all elections."

New Zealand: 1870. " Verbal voting lasted until 1870, when Parliament finally agreed to adopt the secret ballot. Each voter was given a printed ballot paper listing the candidates in their electorate. They marked the paper in private behind a screen and then deposited it into a locked ballot box" (www.elections.org.nz/democracy/history/years.html). Mackie (2000) confirms this year. Przeworski (2010) dates the secret ballot to 1937. The reason for this is that Maoris were not granted the right to a secret ballot until then.

United Kingdom: 1872. This was the year of the Ballot Act, which introduced the Australian Ballot (Asquith, 1888; Mackie, 2000, Seymour, 1915).

Switzerland: 1872. Hewitt (1977) and Engerman (2003).
Canada: 1874. Pillon (2006) and Engerman (2003).
Belgium: 1877. Seymour and Fray (1918, vol. II, p. 193) and Mackie (2000). Przeworski (2010) dates it to 1879, but this is the first election in which it applied.
Norway: 1884. Nerbørvik (1986). Some sources, e.g., Engerman (2003) and Przeworski (2010), give 1885, but this is the year of the first election conducted with secret ballot.
United States: 1891. The dating of the secret ballot for senatorial elections in the USA is complicated by the fact that the ballot rules were decided at the state level. We use 1891 as the benchmark because the majority of states had adopted the secret ballot for senatorial elections in that year (Mackie, 2000).
Denmark: 1901. Elkit (1988, p. 300) writes that secret ballot was finally approved in January and Seymour and Frary (1918, vol. II, p. 177) agree. In fact, the act was passed on 25th january, 1901 by the upper chamber (landstinget). It had already been approved by the lower chamber on 15th November, 1900. However, since final approval from the upper chamber was needed for the act to be passed, we use 1901 as the year of introduction (see Tilæg C til Rigsdagstidenden, 1901).
Finland: 1906. " The Parliament Act that came into force on 1 October 1906 was a monumental reform [...] The new Parliament Act called for Members of Parliament to be elected directly and by secret ballot according to a proportional system based on districts." (http://web.eduskunta.fi/Resource.phx/parliament /aboutparliament/presentation/history.htx). Mackie (2000) and Przeworski (2010) give 1907, but this is the first election under the new rules.
Sweden: 1907. Carstairs (1980). Esaiasson (1990) gives 1911 but this is the first election with secret ballot. Przeworski (2010) dates it to 1866. While it is true that the older and varying regulations often demanded the use of "sealed tickets" (slutna sedlar), it was not till the election in 1911 that voting became de facto secret.
Austria: 1907. (http://www.parlament.gv.at/) and Seymour and Frary (1918, vol. II, pp. 62-63).
Germany: 1913. Anderson (2000, p. 56) argues that the secret ballot was effective in Germany from 1913. Przeworski (2010) dates it to 1867 (i.e., from the constitution of the Northern German Confederation). The constitution of Imperial German, Article 20, guaranteed a secret ballot, but it was not Australian. The ballot papers were not printed centrally but by the candidates themselves. This made it possible by using different types of paper or by putting a mark on the outside of the ballot to monitor the votes. In 1903 a system of "sealed tickets" whereby the votes were concealed in envelopes and the vote became partially secret (Ziblatt, 2009, p. 12). As pointed out by Anderson (1993, p. 1457) and Mares (2015) this, however, did not in many cases make the situation better as the return officer would collect the envelops in the order in which they were cast and in small constituencies it was therefore a simple matter to correlate the sequence of votes with the sequence of voters. Moreover, Anderson (2000, p. 55) notes that " Experts in European elections processes claimed that the ballot envelope actually worsened the chances for secrecy" . In line with this, Ziblatt (2009, p. 17) observes that pettitions claimed violations of the secret ballot in both the 1907 and 1912 elections. In 1913, a system based on mandatory electoral urns removed this loophole and we follow Anderson (2000, 56) and use 1913 to date when the ballot became de facto secret. It is, however, a matter of judgement if 1913 or 1903 is chosen and we note that the last election in Imperial Germany took place in 1912. We have checked that none of the results depends our choice of 1913 and included a discussion of this in appendix S4.
France: 1913. France had semi-secret elections early in the 19th century, but it was not until 1913 that it became effectively secret, e.g. Baland and Robinson (2008, p. 1738), Seymour and Frary (1918, vol. I, p. 379), Markoff (1999) and Crook and Crook (2007). The

constitution of 1795 included provisions for the secret ballot (and Przeworski (2010) dates it to 1820), but it is widely seen to have been ineffective because voters could write the name of their preferred candidate on their own ballot paper at home or receive a ballot in a distribution in the streets. As stressed by many authors (e.g. Seymour and Frary, 1918; Markoff, 1999; Mackie, 2000; Crook and Crook, 2007), this provided ample leeway for corruption and intimidation. For example, Seymour and Frary (1918, p. 379) note that the vote was *de jure* secret, but " in practice almost as public as in Prussia, where it is oral." In 1913, the ballot rules were tidied up and although the ballot remained non-Australian, the reform is widely considered to have been effective in providing secrecy and weeding out most corrupt practices.

US states: Heckelman (1995), Lott and Kenny (1999), Fredman (1968), and Ludington (1911). In some states, e.g., Kentucky, Maryland, Minnesota, Tennessee, Texas, Missouri, and Wisconsin, the secret ballot was not initially applied uniformly throughout the entire state and the coding of these cases follows Lott and Kenny (1999) by coding years of statewide adoption of the Australian ballot. For example, Ludington (1911) notes that Kentucky enacted an Australian ballot law in 1888, but this only applied only to " the election of certain city officers in the city of Louisville" . The first statewide Australian ballot law was enacted in 1892, and we use this date for Kentucky, though this choice make little difference for our results. Based on Ludington (1911), we make some corrections to the dates reported by Lott and Kenny (1999). First, Kentucky is coded as 1882, which is likely to be a typo as the year of statewide adoption was 1892. For Wisconsin, Lott and Kenny (1999) give 1894, but Luddington (1911, p. 78) says 1893.

Tables S1 and S2 contain relevant institutional information for the two samples.

Appendix S3: Definitions and Sources

Sample: Western Europe plus off-shoots

real GDP per capita is real GDP in international 1990 Geary-Khamis dollars, adjusted for the impact of border changes, per capita. Source: Maddison (2003).

urbanization rate is the percentage of the population living in towns with more than 20,000 inhabitants. Missing values are interpolated linearly. Source: Banks (2003) and Mitchell (2003a,b).

electorate/adult population is the electorate (for parliamentary/house elections) in percentage of the adult population. Data are only recorded in election years and assumed to be constant between elections. Suffrage is coded zero for periods without elections. Source: Flora et al. (1983), Mackie and Rose (1991), Mitchell (2003a,b), Cook and Paxton (1998), www.elections.org.nz, and www.elections.ca.

gini coefficient is the Gini coefficient for income inequality. A value of zero (one) expresses total equality (inequality). Data are available with 20-year intervals and missing observations interpolated linearly. Source: Bourguignon and Morrisson (2001, 2002).

learning is defined as a distance weighted average of secret ballot adoptions in other countries:

$$learning_{it} = \sum_j \left(1 - \sqrt{\frac{15 - \#common_{ij}}{15}}\right) A_j(t),$$

where $A_j(t)$ is 1 if country j has adopted the secret ballot at time t and zero otherwise. The variable $\#common_{ij}$ is the number of common notes in the linguistic tree between country pair i and j with the maximum number of common notes being 15. Source: Fearon (2003) and own calculations.

population is the total population in 1000s. Source: Mitchell (2003a,b) and Maddison (2003).

Instrumental variables

revolutionary threat is a weighted sum of revolutionary events taking place in other countries. The weights are the distance between the capitals of each pair of countries. Source: Aidt and Jensen (2014) and Tilly (1993).

distance weighted GDP is a weighted sum of log GDP in other countries where the weights are the inverse distance (in kilometers) between the capitals of each pair of countries. Source: Maddison (2003) and own calculations.

Variables used in robustness checks

Districts (MPs) is the number of MPs elected to the main national parliament and is used as a proxy for the number of districts. Source: Flora et al. (1983), Mackie and Rose (1991), Cook and Paxton (1998).

Voters per MP is the number of voters per MP elected to the main national parliament and is used as an alternative measure of the size of the electorate. Source: Flora et al. (1983), Mackie and Rose (1991), Cook and Paxton (1998).

Effective parties is a dummy variable coded one in the year (and in subsequent years) in which political parties with a national organizational structure started to contest elections and zero before that. Source: Mackie and Rose (1991).

POLITY IV index is an index of democratic institutions, ranging from -10 (extremely autocratic) to +10 (fully functional democracy). Source: Marshall and Jaggers (2000).

Left-wing parties is the share of seats held by left-wing parties in the main national parliament. Source: Mackie and Rose (1991).

Sample: US states

real income per worker is real state output (until 1920) or income (from 1929) per worker in 2000 dollars. Source: Turner et al. (2007).

average years of schooling is the average years of schooling of the workforce, estimated using the perpetual inventory method. Source: Turner et al. (2007).

urbanization rate is the share of the population living in urban areas. Available for census years only. Interpolated linearly for the years in between. Source: Lee et al. (1957) and various US Census reports.

women's suffrage is a dummy taking the value one if women had the right to vote and zero otherwise. Source: Lott and Kenny (1999).

poll tax is a dummy equal to zero in years without a poll tax requirement and equal to one otherwise. Source: Lott and Kenny (1999).

literacy test is a dummy equal to zero in years without a literacy test requirement and equal to one otherwise. Source: Lott and Kenny (1999).

land inequality is the share of land held by the 20 percent largest farms. Source: Galor et al. (2009).

learning is defined as

$$learning_{it} = \sum_j \frac{A_j(t)}{D_{ij}},$$

where D_{ij} is the distance (in miles) between the state capitals of state i and j and $A_j(t)$ is one if state j has adopted the secret ballot at time t and zero otherwise. Source: Own calculations.

population is the number of inhabitants in the state in 1000s. Available for census years only. Interpolated linearly for the years in between. Source: Lee et al. (1957) and various US Census reports.

weighted income per worker is a weighted sum of log *income* in other states. The weights are the distance between state capitals (in miles). Source: Turner et al. (2007) and own calculations.

regional dummies are coded according to the eight Bureau of Economic Analysis regions. Source: www.bea.gov.

turnout rate is equal to the total number of votes cast in a gubernatorial election divided by the age and sex eligible population. Source: Burnham et al. (1971).

Sample: Great Britain

turnout is the total number of votes cast in each one-seat constituency in England, Wales, and Scotland in 1868 and 1874. Source: Craig (1977) and Berlinski and Dewan (2011).

density is inhabitants per house in each constituency in 1861 and 1871. Source: 1861 and 1871 Census of Great Britain.

electorate is the number of registered voters in each constituency in 1868 and 1874. Source: Craig (1977) and Berlinski and Dewan (2011).

Appendix S4: Empirical specification and robustness checks

Duration dependency

We capture duration dependence in the hazard rate through the function $H(.)$. The argument of the function is $t - t_i^p$ where t_i^p represents either the year in which country i enters the " risk set" . To model duration dependence, we estimate $H(.)$ using natural cubic splines and use the estimated spline coefficients along with the number of years a country has been at " risk" of adopting (or since entry to the sample). We use a specification with two knots for the splines.

In the Western Europe plus off-shoots, a formal test for duration dependency in the hazard rate cannot reject that the baseline hazard is constant over time. Our strong prior is that the hazard is increasing with time and we report specifications with duration dependence although it makes no difference to the results.

Western Europe plus off-shoots

First, as mentioned in the main text, we carry out instrumental variables estimations. We instrument for *real GDP per capita* and *electorate/adult population*. As in Acemoglu et al. (2008), we use a weighted index of real GDP in the other countries in the sample as an instrument for GDP per capita in a particular country. The logic is the international transmission of business cycle shocks. Specifically, Acemoglu et al. (2008) calculate their instrument as

$$Y_{it-1} = \sum_{j=1, j\neq i}^{N} \omega_{ij} Y_{jt-1} + \varepsilon_{it-1}$$

where Y is (log) real GDP, N is the number of countries and ω_{ij} is the weight given to real GDP in country j. We use a similar instrument, though the weights (the omegas) are not

defined by trade as in Acemoglu et al. (2008), but simply by the inverse of the distance between two countries as would be warranted by a gravity approach to international linkages. The validity of the instrument can, however, be challenged if social learning effects are strong. This is not the case in our data, but by controlling for the variable *learning*, we can rule out that movements in real GDP in other countries affect the probability of a secret ballot reform in a particular country, not through its effect on real GDP in that country, but through a social diffusion channel. While this instrument, in principle, is relevant for the entire sample, the second instrument only makes sense for the Western European countries. For this reason, the instrumental variable (IV) estimations are restricted to this sub-sample. Aidt and Jensen (2014) demonstrate that revolutionary events (as defined by Tilly (1993)) in other countries affect suffrage reforms in a particular country through a process of international diffusion of information. Revolutionary pressures are unlikely to be a direct cause of ballot reform, and so we use a measure of distance weighted revolutionary events, *revolutionary threat*, in other countries as an instrument for suffrage reform in a particular country. In addition to this, we exploit the high degree of path dependency in suffrage rules and make use of the one-year lag of *electorate/adult population* as an instrument. This allows us to test the validity of the extra instrument using an over-identification test.

Table S4 reports the IV results. The IV estimates are based on a linear probability model, and for the smaller sample without the off-shoots. For comparison, we report the results from a logit and a linear probability model on this smaller sample in columns one and two. The IV estimates in column three confirm the modernization hypothesis and reject the franchise hypothesis. The first stage regressions are reported in the last two columns. The instruments are highly significant with large F-statistics for joint significance and pass the J-test for over-identification. This suggests that the correlation between modernization and the timing of the secret ballot does, in fact, represent a causal mechanism.

Second, we have, based on information from Flora et al. (1983), Mackie and Rose (1991), Cook and Paxton (1998), constructed an alternative measure of the size of the electorate, *voters per MP*, defined as the number of voters per seat of parliament. This measure is also insignificant. Importantly, the results for the modernization variables are unchanged.

Third, we have estimated all models with a random effects logit estimator. The test of country specific random effects fails to reject the null of no country specific effects. Importantly, the modernization variables remain significant.

Fourth, we measure landholding equality by the variable *share of family owned farms* (Vanhanen, 2003). This variable is only available from 1858. Accordingly, by including it in the model, we lose more than half the observations and three countries. The variable is not significant. The coefficient on real GDP per capita is significant and the two modernization variables remain jointly significant. The coefficient on *electorate/adult population* is negative but insignificant.

Fifth, we have checked that the results are unaffected if we, instead, date the secret ballot in the USA to 1896 (rather than 1891) when ninety percent of the states were using the secret ballot (Mackie, 2000). We have also checked that the results are robust to excluding the USA or all the off-shoots from the estimations.

Sixth, one can question whether *population* belongs to a model in which we are interested in estimating the effect of real GDP per capita and other variables measured relative to population size. The argument in favor is that scale effects could be important and by not including population directly, we could introduce omitted variables bias. The argument against is that by keeping the denominator of all the ratio variables constant, we estimate the various effects from variation in the nominator. On balance, we decided to report specifications with population size included. We have re-estimated all specifications without

population. It makes no difference to the results.

Seventh, as noted in appendix S2, we have coded the adoption of the secret ballot in Germany to be 1913. We have checked that using the alternative coding (1903) makes no difference to the results.

Eighth, in the main specification, we control for the extension of the suffrage but other institutional aspects could also be important. In Table S5, we investigate two suchs aspects. First, to control for the number of districts (and the fact that ballot reforms sometimes conincide with redistricting), we add the variable *district (MPs)* to the baseline. We see that the variable is negatively correlated with the adoption probability. Second, we add the *POLITY IV index*. This index is positively correlated with the adoption probability, but insignificant. In both cases, the evidence on the modernization effect remains strong.

Ninth, we investigate the possibility of a nonlinear relationship between the *real GDP per capita* and the adoption probability. In Table S5, we show that there is some evidence of a U-shaped relationship for this sample but we cannot replicate this in the US state sample [not reported].

Finally, we observe that the first principal component is strongly and positively correlated with both modernization variables with a correlation coefficient of nearly 0.93. Thus, this variable captures modernization much better than the second principal component which is negatively correlated (correlation coefficient=-0.35) with urbanization rate and positively correlated with *GDP per capita* (correlation coefficient =0.35).

The US states: the event history study.

In Table S6, we present estimations in which we instrument for *real income per worker*, *average years of schooling*, and *urbanization rate*. We use a weighted index of real income in other states along with regional dummies for the eight economic areas of the USA as instruments. These capture that states that are located in the same region share similar geographical conditions, e.g., access to the sea, which are likely to affect economic development (GDP, urbanization, and human capital accumulation) but not, conditional on *learning*, the timing of the secret ballot. For comparison, we report the results from a logit and a linear probability model in columns one and two. The IV estimates in column three shows that the positive effect of *average years of schooling* is robust to instrumentation and that the three instrumented modernization variables are jointly significant. The first stage regressions are reported in the last three columns. The instruments are strong and pass the J-test for over-identification. This points to a causal relationship running from modernization to secret ballot.

Boix (2003, p. 122) notes that racial motives might have played a role in relation to the secret ballot. The share of blacks in the population is itself insignificant and it has no effect on our main results. We have re-estimated the model for the period after the civil war and the 15th amendment (1870 onwards) and without population size. Doing so, again, matters little for the results. Adding state specific random effects also has little effect on the results. The test of heterogeneous random effects across states fails to reject the null of no state specific effects. The three modernization variables correlate positively with the first component (correlation coefficients above 0.87). As seen in Table 2, column 5, the first component is highly significant with a positive sign. As in the case of Western Europe and off-springs, the other components do not measure modernization well.

We have investigated the possibility of a nonlinear relationship between the *real income per worker* and the adopt probability and find that the relationship has overall significance, but the individual coefficients are imprecisely estimated.

The factor loading are 0.9291 (for income), 0.9063 (for schooling) and 0.8703 (for

urbanization).

The US states: the turnout model.

The outcome variable (*turnout rate*) is a fractional variable bounded between zero and one. Papke and Wooldridge (1993) propose to use a logit link for fractional variables instead of the linear estimator. We find that using this method matters little for the results. We have added a linear trend to all the specifications. The trend is not significant and it makes no difference to the variables of interest. We include women's suffrage in the regression model. This raises the question of time. Some states introduced women's suffrage before the secret ballot (e.g., Wyoming, Utah, Georgia, Tennessee, North and South Carolina) while others did in 1920 when it became mandatory to do so and in many cases years after the secret ballot was introduced. We have re-estimated the panel model with women's suffrage and the results regarding modernization are unaffected. Another issue is that while many states adopt the ballot between 1888 and 1896, there are a few which adopt much later (South Carolina in 1950 and North Carolina in 1929). We have checked that these 'temporal outliers' do not influence the results by re-estimating the model without them. The results of interest are unaffected.

Great Britain

Our sample comprises 65 borough constituencies in England and Wales and 9 county constituencies in Wales and Scotland. This is not the entire universe of one-seat constituencies because data are sometimes missing and the 1868 or 1874 elections were sometimes uncontested. The results are robust to controlling for population growth (which in itself is not significant).

Supplementary references

Acemoglu, Daron., Robinson, James A., 2000. Why did the west extend the franchise? Democracy, inequality, and growth in historical perspective. Quarterly Journal of Economics 115(4), 1167-1199.

Anderson, Margaret L., 1993. Voter, Junker, Landrat, Priest: The old authorities and the new franchise in Imperial Germany. The American Historical Review 98(5), 1448-1474.

Anderson, Margaret L., 2000. Practicing democracy. Elections and political Culture in Imperial Germany. Princeton Universit Press, Princenton, New Jersey..

Asquith, H.H., 1888. The ballot in England. Political Science Quarterly 3(4), 654-681.

Banks, Arthur S., 2003. Cross-national Time Series 1815-2003. ICPSR (ed.). Inter-university Consortium for Political and Social Research, Ann Arbor, MI.

Bourguignon, Francois. and Morrisson, Christian, 2001. Data sources for inequality among world citizens, 1820-1992 [computer file]. Unpublished, available at www.delta.ens.fr/XIX.

Bourguignon, Francois, Morrisson, Christian, 2002. Inequality among world citizens 1820-1992. American Economic Review 92(4), 727-744.

Brewer-Carias, Allan-Randolph, 1997. Las constituciones de Venezuela. Academia de Ciencias Polticas y Sociales: Caracas, Venezuela.

Burnham, W. Dean, Clubb, Jerome M., Flanigan, William, 1971. State level congressional, gubernatorial and senatorial election data for the United States, 1824-1897. ICSPR 0075.

Cook, Chris, Paxton, John, 1998. European political facts. MacMillan Press, London.

Crook, Malcom, Crook, Tom, 2007. The advent of the secret ballot in Britain and France, 1789–1914: From public assembly to private compartment, History 92(308), 449-471.

Elklit, Jørgen, 1988. Fra å ben til hemmelig afsteming. Politica, Aarhus, Denmark.

Engerman, Stanley L., 2003. The history and political economy of international labour standards. In Kaushik Basu, Henrik Horn, Lisa Roman, Judith Shapiro (eds.): International labour standards: history, theory, and policy options, Blackwell Publishing.

Esaiasson, Peter, 1990. Svenska valkampanjer, 1866-1988. Almänna Förlaget, Stockholm.

Flora, Peter, with Jens Alber, Richard Eichenberg, Jürgen Kohl, Franz Kraus, Winfried Pfenning, and Kurt Seebohm, 1983. State, Economy and Society 1815-1975, Vol. I. Campus Verlag, Frankfurt, Germany.

Hewitt, Christopher, 1977. The effect of political democracy and social democracy on equality in industrial societies: A cross-national comparison, American Sociological Review 42(3), 450-464.

Lee, Everett S., Miller, Ann R, Brainerd, Carol P, Easterlin, Richard A. (eds.), 1957. Population redistribution and economic growth: United States, 1870–1950. Vol. 1. American Philosophical Society, Philadelphia.

Ludington, Arthur C., 1911. American ballot laws, 1881-1910. Bibliolife, UK.

Mackie, Tom T., Rose, Richard, 1991. The international almanac of electoral history. Macmillan, London.

Maddison, Angus, 2003. The world economy: historical statistics. Paris: OECD.

Markoff, John, 1999. Where and when was democracy invented? Comparative Studies in Society and History 41(4), 660-690.

Marshall, M.G., Jaggers K., 2000. Polity IV Project. Data set Users Manual. Center for International Development and Conflict Management, University of Maryland

Mitchell, Brian R, 2003a. International historical statistics: Asia & Oceania, 1750–2000, 4th ed. Palgrave Macmillan, Basingstoke.

Mitchell, Brian R,, 2003b. International historical statistics: the Americas, 1750–2000, 5th ed. Palgrave Macmillan, Basingstoke.

Nerbørvik, Jostein, 1986. Norsk historie, 1870-1905: Fråjordbrukssamfunnmot organisasjonssamfunn, Oslo, Det Norske Samlaget.

Papke, Leslie E., Wooldridge, Jeffrey M., 1996. Econometric methods for fractional response variables with an application to 401(k) plan participation rates. Journal of Applied Econometrics 11(6), 619-632.

Pillon, Dennis, 2006. Explaining voting system reform in Canada, 1874 to 1960. Journal of Canadian Studies 40(3), 135-161.

Stuurman, Siep, 1991. 1848: Revolutionary reforms in the Netherlands. European History Quarterly 21(4), 445-480.

Tillæg C til Rigsdagstidenden, 1901. Vedtagne lovforslag, beslutninger m.m. 53de ordentlige samling 1900-1901. Copenhagen, T.B. Schultz.

Turner, Chad, Tamura, Robert, Mulholland, Sean, Baier, Scott, 2007. Education and income of the states of the United States: 1840-2000, Journal of Economic Growth 12(2), 101-158.

Vanhanen, Tatu, 2003. Democratization: A comparative analysis of 170 Countries. Routledge, London.

Table S1: Institutional information for the Western Europe plus off-shoots sample.

Country (year of entry to sample)	Year of adoption of the secret ballot	Year of franchise extensions (1820-1913)	*Electorate/adult population* in year of first election or at independence (in year of adoption)	Year of establishment of parties with nation-wide organizational structures
Western Europe				
Netherlands (1830)	1849	1848, 1887, 1894	4.6 (4.6)	1880
Switzerland (1848)	1872	1848	38.9 (38.9)	1890
United Kingdom (1820)	1872	1832, 1867, 1884	8.6 (14.9)	1880
Belgium (1830)	1877	1831, 1848, 1893	1.9 (3.7)	1846
Norway (1820)	1884	1814, 1884, 1897	11.4 (11.4)	1882
Denmark (1820)	1901	1849	25.7 (29.0)	1870
Finland (1820)	1906	1869, 1906	8.3 (76.2)	1890
Austria (1867)	1907	1867, 1896, 1907	10.6 (37.7)	1889
Sweden (1820)	1907	1866, 1907	9.8 (14.0)	1887
France (1820)	1913	1824, 1830, 1848	0.5 (43.4)	1893
Germany (1871)	1913	1871, 1919	33.0 (38.7)	1871
Western off-shoots				
New Zealand (1856)	1870	1860, 1867, 1879	29.1 (33.8)	1890
Canada (1867)	1874	1898	22.0 (23.1)	1867
USA (1820)	1891[a]	1870	n.a. (40.6)	1828

Notes: Italy is not included in the sample because it *de facto* adopted the secret vote in 1861 at unification. Australia is not included in the sample because the secret ballot was introduced at the time of independence. a. This is the year when the majority of the US states has Australian ballot.
Sources: See appendix S2.

Table S2: Institutional information for the US states sample.

State	Australian Ballot	Women's suffrage	Poll tax	Literacy test
Massachusetts	1888	1920	-1891	1857-
Indiana	1889	1919		
Montana	1889	1914		
Rhode Island	1889	1917	-1888	
Mississippi	1890	1920	1889-1963	1890-
Oklahoma	1890	1918		1912-
Vermont	1890	1920		
Washington	1890	1910		1896-
Wyoming	1890	1869		1889-
Arizona	1891	1912		1912-
Arkansas	1891	1917	1891-1963	
California	1891	1911		1894-
Colorado	1891	1893		
Delaware	1891	1920	-1907	1897-
Idaho	1891	1896		
Illinois	1891	1913		
Maine	1891	1919		1892
Michigan	1891	1918		
Minnesota	1891	1919		
Missouri	1891	1919		
Nebraska	1891	1917		
Nevada	1891	1914	-1910	
New Hampshire	1891	1920		1902-
North Dakota	1891	1917		
Ohio	1891	1919		
Oregon	1891	1912		1924-
Pennsylvania	1891	1920	-1933	
South Dakota	1891	1918		
West Virginia	1891	1920		
Iowa	1892	1919		
Kentucky	1892	1920		
Maryland	1892	1920		
Alabama	1893	1920	1901-1963	1901-
Kansas	1893	1912		
Wisconsin	1893	1919		
Virginia	1894	1920	1875-1882, 1902-1963	1902-
Florida	1895	1920	1889-1927	
New York	1895	1917		1921-
Louisiana	1896	1920	1898-1934	1898-
Utah	1896	1870		
Texas	1905	1918	1902-1963	
Connecticut	1909	1920		1856-
New Jersey	1911	1920		
New Mexico	1912	1920		
Tennessee	1921	1919	1870, 1890-1951	
Georgia	1922	1920	-1945	1908-
North Carolina	1929	1920	1899-1920	1900-
South Carolina	1950	1920	1895-1951	1895-

Sources: Heckelman (1995), Lott and Kenny (1999, table 1), and Ludington (1911).

Table S3: Descriptive statistics for the variables used in the estimations.

Variable	Obs.	Mean	Std. dev.	Min	Max
Western Europe plus off-shoots					
Real GDP per capita	2681	4753.56	4541.49	400	28129.23
Urbanization rate	1488	197.45	135.91	0	631
Electorate/adult population	1732	31.48	31.01	0	101.90
Gini coefficient	1403	0.52	0.02	0.48	0.56
Learning	1091	0.16	0.25	0	1.32
Population (in 1000s)	2722	24571	38601	70	290343
Revolutionary threat	1452	0.0003	0.001	0	0.015
Distance weighted GDP	1872	0.35	0.10	0.12	0.56
US states					
Real income per worker	4778	12043.2	6336.738	1989.977	47727.45
Average years of schooling	4798	4.92	2.55	0.20	11.11
Urbanization rate	4828	32.29	21.84	0	92.62391
Share of land held by the 20%	1225	0.16	0.053	0.006	0.47
Learning	2627	0.008	0.02	0	0.10
Population	4858	1,682,579	1,942,818	6,077	14,900,000
Distance weighted income	5220	1.64	0.58	0.28	2.89
Turnout rate	1426	52.30	19.30	2.13	99.99
Great Britain					
Turnout	74	2260	1812	572	8496
Density	74	5.6	1.8	4	13.5
Electorate	74	2813	2391	706	10352

Note: We do not report statistics for dummy variables.

Table S4: The effect of modernization on the probability of secret ballot in the Western Europe sample. Instrumental variables estimations.

Model	(1)	(2)	(3)	(4)	(5)
	Reform	Reform	Reform	Log(GDP per capita)	Electorate/ adult population
log(real GDP per capita)	8.41*	0.064**	0.065**		
	[1.84]	[3.08]	[2.33]		
Electorate/adult population	0.047	0.00065	-0.00015		
	[1.27]	[0.81]	[-0.47]		
Instrumental variables					
Distance weighted GDP				2.09***	1.28
				[26.49]	[0.66]
Lagged electorate/adult population				-0.003***	0.98***
				[-6.70]	[77.60]
Revolutionary threat				0.0067*	0.35***
				[1.66]	[3.55]
Control variables					
Gini coefficient	9.76	0.086	0.15	4.10***	3.86
	[0.29]	[0.33]	[0.51]	[10.37]	[0.40]
log(population)	-0.90**	-0.011**	-0.0083**	0.099***	0.12
	[-2.46]	[-2.45]	[-1.99]	[18.21]	[0.91]
Learning	-2.08	-0.022	-0.0043	0.18***	-0.39
	[-1.32]	[-0.81]	[-0.14]	[6.22]	[-0.538]
Control for duration dependence and constant	Yes	Yes	Yes	Yes	Yes
Observations	692	692	688	688	688
First stage F-test				236.36***	2227.99***
J-test (over-ID test)			0.686		
Number of countries	11	11	11	11	11
Estimation method	Logit	OLS	2SLS	First stage	First Stage

Notes: See Table 1. The off-shoots are not included in the sample.

Table S5: Selected additional robustness checks for Western Europe plus off-shoots sample. Dependent variable: *reform*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(real *GDP* per capita)	6.36**	6.36**	7.054*	6.47**	10.15***	7.18**	-105.3**
	[2.44]	[2.41]	[1.73]	[2.50]	[3.14]	[2.45]	[2.00]
log(real *GDP* per capita), squared							7.39**
							[2.05]
Electorate/adult population	0.037	0.037	-0.0135	0.037	0.04	0.04	-0.11*
	[1.26]	[1.25]	[-0.47]	[1.32]	[1.55]	[1.44]	[1.83]
Electorate/adult population, squared							0.003**
							[2.60]
Gini coefficient	28.49	28.59	21.72	30.31	37.9	25.79	44.92
	[1.44]	[1.39]	[0.80]	[1.40]	[1.28]	[1.45]	[1.63]
log(population)	-0.50**	-0.50*	-0.40	-0.50**	-0.42	-0.104	-0.72***
	[-2.04]	[1.67]	[1.60]	[2.18]	[1.86]	[0.36]	[2.89]
Learning	-1.18	-1.18	0.456	-1.54	-0.78	-1.72	-0.23
	[-0.94]	[-0.92]	[0.293]	[1.20]	[0.52]	[1.17]	[0.27]
Left-wing parties		0.0003					
		[0.007]					
Polity IV			0.05				
			[0.62]				
Effective parties				0.55	49.8		
				[0.49]	[1.50]		
Effective parties*log (real GDP per capita)					-6.33		
					[1.50]		
Districts (MPs)						-0.004*	
						[1.65]	
Control for duration dependence	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	779	779	691	779	779	779	779
Estimation method	Logit	Logit	Logit	Logit	Logit	Logit	Logit

Notes: Robust z-statistics correcting for clustering by country in brackets (i.e., we allow the errors to be correlated over time within countries, but not between countries). Constants not reported. All estimations based on sample of 14 countries. *** p<0.01, ** p<0.05, * p<0.1.

Table S6: The effect of modernization of the probability of secret ballot in US states, 1840-1950. Instrumental variables estimation.

	Austra-lian ballot	Austra-lian ballot	Austra-lian ballot	Log(real income per worker)	average years of schooling	Urbanize-tion rate
The modernization hypothesis						
log(real income per worker)	1.239**	0.010**	0.0165			
	[2.374]	[2.12]	[1.459]			
average years of schooling	0.324*	0.0027	0.0101*			
	[1.910]	[1.11]	[1.957]			
urbanization rate	-0.00661	0.000465	-0.00059			
	[-0.384]	[1.30]	[-1.127]			
Joint significance of modernization variables	**20.74*** **	**7.68*** **	**37.14*** **			
The franchise hypothesis						
women's suffrage	0.156	0.00841	0.015	-0.264*	-0.088	0.569
	[0.331]	[0.394]	[0.706]	[-1.935]	[-0.514]	[0.261]
literacy test	-0.775	-0.0402*	-0.00325	0.123	-0.331**	4.536
	[-1.086]	[-1.86]	[-1.306]	[1.153]	[-2.047]	[0.934]
poll taxes	-0.254	-0.014*	-0.001	-0.021	-0.586**	3.691
	[-0.614]	[-1.246]	[-0.0772]	[0.261]	[-2.034]	[0.47]
Joint significance of franchise variables	**1.72**	**1.73**	**4.16**			
Control variables						
log(population)	-0.0877	-0.00084	0.00195	-0.0022	0.19***	3.21*
	[-1.027]	[-0.45]	[0.676]	[-0.082]	[2.72]	[1.91]
Learning	-5.353	0.64	0.75	-0.0209	3.08	69.35
	[-0.385]	[1.27]	[1.61]	[-0.0307]	[0.80]	[0.94]
Instrumental variables						
Weighted income per worker				0.0215	0.00096	10.48
				[0.136]	[0.0414]	[1.51]
New England				-0.423	0.828**	-7.489
				[-1.508]	[2.21]	[0.719]
Mid East				-0.343	-0.636	-7.602
				[-1.209]	[-1.512]	[-0.826]
Great Lakes				-0.436*	-0.394	-20.02***
				[-1.876]	[-0.97]	[3.535]
Plains				-0.524**	-0.77***	-18.51***
				[-2.412]	[-2.135]	[-4.442]
South East				-1.065***	-2.265***	-28.65***
				[-4.975]	[-7.146]	[-4.939]
South West				-0.063***	-2.329	-19.53***
				[-3.375]	[-4.051]	[-4.664]
Far West				0.216	-0.179	-2.583*
				[1.030]	[0.638]	[0.506]
First stage F-test				10.34**	44.00**	10.46**
J-test (over-ID test)			3.94			
Estimation method	Logit	OLS	2SLS	First stage	First stage	First stage
Control for duration dependence and constant	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,186	2,186	2,186	2,186	2,186	2,186

Notes: See Table 2. Excluded region is Rocky Mountains. The number of states is 44.