European
Historical
Economics
Society

EHES Working Paper | No. 288 | November 2025

Resisting reform in the public sector: theory and historical evidence from the UK

Oliver Brufal[1]
University of Groningen

https://www.ehes.org/

EHES Working Paper | No. 288 | November 2025

Resisting reform in the public sector: theory and historical evidence from the UK

Oliver Brufal[1]
University of Groningen

Abstract

Reform in the public sector treats organisations as equal under one law. This is despite large inequalities in practical, de facto, powers. I argue that these inequalities can distort meritocratic reforms. I collected and digitised individual-level data on thousands of entrants to the British Civil Service from 1864-1875. Exploiting an 1870 reform that forced meritocratic hiring on a large part of, but not all, the public sector, I find that patronage use declined among treated organisations, but was not extinguished in spite of comprehensive legislation. I find that variation in non-compliance among treated organisations is driven by heterogeneity in de facto powers. Powerful organisations continue to use patronage after reform due to their exploitation of 'grey areas' in legislative design. My analysis shows a clear issue with assuming external validity from existing work that focuses on lower status, frontline, public sector organisations; the tail does not wag the dog.

JEL Codes: N43, D73, M51
Keywords: Britain, public sector organisations, institutional change, meritocracy

[1] Corresponding Author: Oliver Brufal (o.l.r.brufal.de.melgarejo@rug.nl)
I thank Thilo Albers, Leticia Arroyo Abad, Jutta Bolt, Agnes Cornell, Giampaolo Lecce, Mikolaj Malinowksi, Johann Ohler for helpful comments. I further thank seminar and workshop participants at Leiden, Groningen,Munster, and Oxford for valuable suggestions. I am indebted to the staff at the National Archives at Kew for their help during data collection.

Notice
The material presented in the EHES Working Paper Series is property of the author(s) and should be quoted as such. The views expressed in this Paper are those of the author(s) and do not necessarily represent the views of the EHES or its members.

1 Introduction

Bureaucracy is a crucial component of state capacity and economic development (Besley, Burgess, et al. 2022). Efforts to improve the quality of bureaucracy are multi-faceted (World Bank 2000), yet a crucial approach is to introduce merit-based selection procedures to replace patronage systems that often prioritise rent-seeking above public service delivery (Robinson and Verdier 2013).[1] Meritocratic systems are conventionally associated with higher levels of state capacity and public good provision (Rauch and Evans 2000). In spite of this, there is limited empirical research into how the transition to merit-based hiring should be structured to prevent the persistence of informal patronage networks after reform.[2] Due to the size of bureaucracy and the issues with quantifying output among non-frontline positions, there is a sensible tendency to focus upon frontline public service delivery among existing research. However, the result is that we know little about how higher-tier organisations react to reform, and thus whether there is true external validity for the idea that meritocratic reform can quickly reach, and improve, all organisations within the cross-section of a national bureaucracy.

The central issue is that there may be differences in power between the organisations that constitute the public sector. By power, one should think of *de facto* rather than *de jure* powers. This is best defined in the public sector as the practical capacity (or powers) an organisation has (in bargaining terms), akin to a hierarchy. Often these are the result of the function an organisation has, or the historical importance it has had to state development. If all these powers remain constrained by the state–i.e. regulatory bodies–then this can be a limited concern. But, in settings where the state lacks the capacity to constrain unwieldy groups (which I refer to as a weak institutional climate hence) I suggest that these inequalities can in turn create heterogeneity in reform success along this same gradient.[3] As such, meritocratic reforms in general equilibrium may not be as immediately consequential as frontline evidence might suggest. As of yet, there has not been any research which assesses the extent of this problem, which this paper attempts to rectify.

In this paper I provide causal empirical evidence on how meritocratic reform affects the bureaucracy as a whole, rather than focusing on individual organisations. This scale allows me

[1] The improvement delivered by merit systems has a long intellectual heritage in the Weberian ideal of an impartial bureaucracy whose lack of vested interests leave public service as the primary concern rather than the narrow interests of a clientelist relationship (Weber 1922). Historically, it has been well noted that bureaucracies operated through patronage, and one of the crucial developments in improving state capacity has been this switch away from narrowly defined interests (Kettering 1988).

[2] The conventional research on patronage in political science has focused on quantifying its consequences (Folke et al. 2011) or to consider the explanations for its reform (Ting et al. 2013; Anzia and Trounstine 2025). A recent paper expands upon this to consider the demographic consequences of civil service reforms across US municipalities (Kuipers and Sahn 2023). The same is true of research in economics, which has focused in historical (Aneja and Xu 2024; Moreira and Pérez 2024) and contemporary (Colonnelli et al. 2020; Estrada 2019) settings on the deleterious effects of patronage rather than how it should be transitioned away from. This evidence without research on implementation provides a clear policy recommendation with no advice on how it should be implemented, which has negative consequences for policymakers looking to follow academic research.

[3] The success of the British economy in the 19th century may lead one to be hesitant to describe it as a weak institutional climate. Famously, institutional economic historians have often predicated its success upon its institutions (North and Weingast 1989). However, it is clear that state institutions were dominated by clientelist networks, and that these networks were able to override policy directives and internal regulation. As such, while Britain was moving towards an increasingly democratic and meritocratic institutional landscape, as of 1870 it could not be considered strong (Xu 2018; Chapman 2025). This is not to dispute that in relative terms British institutions were strong in a cross-country comparison, as it purely judges strength by the capacity to constrain the type of corrupt patronage networks that are the point of study here.

to provide unique evidence on how, and why, non-compliance occurs when organisations have meritocratic reforms imposed on them.[4] I focus on one potential theoretical explanation, variation in de facto power among treated organisations. I derive a brief theoretical framework that formalises the interaction of de facto power and compliance through the incentives of managerial bureaucrats in charge of a single hiring decision. The model assumes that bureaucrats derive utility from exercising patronage and incur no disutility from adhering to meritocratic constraints other than the loss of previous utility gain. Reform creates new costs to exercising patronage in a meritocratic environment, which are incorporated into the model as the probability and consequence of punishment for non-compliance. By maximizing their utility, bureaucrats weigh the benefits of patronage against the costs of punishment. My hypothesis is that de facto power among organisations allows them to bargain with new regulation. As such, high capacity organisations are able to exploit failures in legislative design and use their strength to insulate themselves from public (and politician) oversight and criticism. In outcome terms, more powerful organisations are able to maintain previous selection habits, whereas weaker counterparts are compelled by legislation to adopt merit-hiring. I use a unique empirical setting to test the accuracy of this hypothesis.

My empirical evidence comes from the first meritocratic public service reform implemented in the West, the British Order in Council of 1870.[5] Previous to this reform, the civil service had relied on patronage to fill entry level positions. This had led to public characterisations of the service as befouled by nepotism and inept appointments (Chapman and Greenaway 1980). It was thought that this was corroding the quality of public service delivery over time, although this may be more reflective of increased scrutiny and accountability from the general public as wealth and education spread throughout the 19th century.[6] Despite the importance of this reform, there has been no quantitative analysis of its consequences on personnel practices.[7] I digitised 11 years of data from the *Reports of the Civil Service Commission.* The reports provide the universe of Civil Service entrants at the clerk-level, and given a peculiarity of the registration system (that I return to later) allow me to observe their route of entry (patronage or merit) consistently throughout the period. This individual-level data enables me to construct a unique dataset of nearly 10,000 civil servants. This is the first time the reports have been assembled into a single individual-level dataset.

Empirically, the chosen reform has several empirical properties that make it desirable for my chosen question. The first is the depth of data, which is gathered across the British Civil Service rather than a subset. Second, there is non-uniform reform imposed by the government upon its

[4] Throughout the paper I refer to organisations as 'departments' given that this was the historically-used term in the British Civil Service – given the preference for organisations in the contemporary literature I use the two interchangeably.

[5] This transition to meritocracy has been claimed to have been central to the American Pendleton Act of 1883, as British public sector organisations were used as counterfactuals by US reformers to stress the costs of patronage (Johnson and Libecap 1994a).

[6] The question of causality in why a society adopts civil service reform has been of particular interest to political scientists focused on the Pendleton Act. The conventional answer in this case, and which is likely in others, is that a combination of voter pressure, transaction costs and inter-party competition forced change (Johnson and Libecap 1994b; Theriault 2003; Ting et al. 2013). A recent paper advocates the view that labour movements helped drive reform (Anzia and Trounstine 2025), an area that offers significant potential for future research.

[7] A previous article focuses on the diffusion of meritocracy between the Indian Civil Service and other civil service departments, but not on the dynamics of the reform itself (Cornell and Svensson 2023). Neither has there been considerable research into the (socio-) economic consequences of the reforms either, such as that done in the US (Moreira and Pérez 2022; Aneja and Xu 2024). This is not the focus of this article, but is certainly worthy of future research to see whether conclusions found there are generalisable to Europe at the same time.

public sector organisations. This is rare, particularly in the modern day, as one would typically exercise such a reform unilaterally. This provides a natural control group of untreated organisations. Related to this is the benefit that reform was not anticipated nor selected into at the organisational level, preventing endogeneity issues related to deliberate policy choices (Ujhelyi 2014a). While I deal with the question of selection in more depth later, these issues are avoided because the reform was itself designed as an experiment by the governing Liberal Party. As William Gladstone, Prime Minister at the time, stated in Parliament, the reform was implemented on 'a scale quite sufficient, even if there should be exceptions, to enable the public to test its principle on a perfect scale' (Parliament of the United Kingdom 1870c). As the first Western country to implement meritocratic reform, the presence of a control group was seen as vital to understanding whether it worked. Finally, and perhaps most importantly, due to unique regulations, there is no reason to believe that patronage hires are 'missing' from the records. As such, it is plausible that these records constitute the true universe of entrants, and thus allow identification that in other settings would be clouded by sample bias.

Given these properties, I am able to use a difference-in-differences strategy to causally estimate the effect of reform. This allows me to determine the success of policy through the comparison of reformed and unreformed departments (organisations, referred to as departments henceforth for historical accuracy). By exploiting *within-department* variation I provide a causal average treatment effect on the treated (ATT). This DiD approach also enables me to test for institutional resilience by exploring heterogeneous treatment effects using a triple difference specification.

My empirical analysis yields two major results. First, exploiting within-department variation in selection, I find that *reformed* departments did in fact reduce patronage use relative to the control group of unreformed departments. Reformed departments were 34-40 percentage points less likely to have an patronage entrant than unreformed after 1870. Panel estimates support this, suggesting that the proportion of new hires entering through open competition increased by 33-38 percentage points. This is despite evidence of positive spillovers from the treated to the untreated group which attenuate any estimates, but do not challenge a causal interpretation. This makes clear that top down policy did have immediate positive effects on meritocratic selection in the Civil Service, with event study estimates confirming an immediate and persistent effect of reform. This coefficient, however, does not indicate that selection had become entirely homogenised, as was the stated policy goal. Patronage hires continued, albeit in a reduced manner. Back of the envelope calculations indicate that they made up around 30-50% of all civil service hires annually after reform.

Motivated by my framework, I use a triple difference specification to explain non-compliance at the department level. My proposed hypothesis is that departments with greater status would be better able to hold off reform efforts, and so they would show smaller treatment effects than departments of a lower status.[8] To operationalise this, I create two binary measures of pre-treatment status within the civil service, and incorporate these as the third difference in my model. These distinctions, 'Ancient' and 'Whitehall' are motivated by historical and geographical prominence

[8] Status is a package term for several correlated traits that departments would share, they would also be more powerful in bargaining and resource terms, hold more instrumental roles in governance and perhaps be able to bend the will of regulatory bodies.

within the civil service before treatment occurred to avoid tautological issues. I find that the aggregate results are driven by less powerful departments, who were compelled into change. The most powerful group, the 'Ancient', show no significant difference from the control, and are around 50 percentage points more likely to have a candidate enter through patronage post-reform than non-Ancient treated departments. Using disaggregated data on type of patronage, I show that powerful departments exploited gray areas in the design of legislation to a degree that allowed them to continue as before. This confirms my hypothesis that these departments had the capacity to invest to protect their previous privileges. I rule out alternative mechanisms, as well as providing qualitative evidence from contemporaries to support my preferred interpretation.

These results suggest that when taking a civil service as a cross-section, meritocratic reforms may have uneven effects when regulatory bodies are weak, which is often the case outside of the modern developed world. In such a climate, the most powerful organisations may circumvent reform, while the weak are forced to comply.[9] As such, aggregate success may mask a failure to tackle patronage in the most crucial organs of the state - which has important equity and potential performance consequences. A clear policy solution does not follow cleanly, for success reforms might either specifically target this heterogeneity in power, or adopt the strategy of successive reform followed by the British. That this took until the 1930's to achieve full success (with the Warren Fisher Reforms) might concern those attempting to instigate policies in politically unstable environments. Generally, they indicate that the general equilibrium effects of meritocracy within the public sector should be considered, alongside the most easily quantified in output terms.

Related Literature

The findings here provide clear evidence on meritocratic reform. This area of research has been of heavy interest to both economists, economic historians and political scientists alike. As such, this article is able to provide new evidence, particularly on where meritocratic reforms fail, that can help to inform distinct areas of interest in multiple disciplines.

The first area of contribution is to the literature on patronage in political science (Grindle 2012). Existing research mainly focuses on the consequences of patronage as public loss or narrowly defined (typically electoral) gain (Folke et al. 2011), why merit systems are implemented in their places (Anzia and Trounstine 2025; Ting et al. 2013), the societal consequences to reform (Kuipers and Sahn 2023) and its potential perverse consequences (Ujhelyi 2014a; Ujhelyi 2014b). This research has moved the literature beyond initial descriptive analyses of what patronage looked like and instead considers its broader context (Sorauf 1956).[10] It is notable however that there has been no empirical or theoretical consideration on how patronage reform is implemented. This consideration is a complication for most studies, which seek to employ a patronage/merit binary to guide their research design. Yet, understanding the dynamics of patronage reform, particularly where it is susceptible, provides a complementary avenue for research to dive beyond

[9] In a stronger institutional climate, all organisations remain under the law despite heterogeneous powers. As such one would not expect any deviations from merit in such an environment.

[10] This research typically relied on case studies (Johnston 1979) or loose theoretical frameworks (Wilson 1961) that debated the reasons behind and implementation of patronage.

this (at times helpful) abstraction.[11] My article introduces this avenue to the existing literature.[12] A further contribution within the literature is that almost all existing research is focused on historical patronage in the US, in particular that which considers reform. The peculiarities of the US reform process, particularly the gulf between federal, state and local bureaucracies (Bostashvili and Ujhelyi 2019), make it important to consider transitions elsewhere as a comparison.

Second, my findings speak to a highly complementary literature on personnel economics within state organisations (Finan et al. 2015). This literature typically focuses on observational studies which quantify the public service effects of discretionary (patronage) versus merit appointments (Colonnelli et al. 2020; Xu 2018), or provide experimental evidence on how different adjustments can improve selection and on-the job performance (Ashraf et al. 2020; Deserranno 2019). Patronage systems have been shown to generally impact public service delivery negatively (Estrada 2019).[13] Recent historical papers exploiting the transition to merit-hiring in the now developed world have allowed a rich natural experiment, with the evidence generally confirming this hypothesis (Aneja and Xu 2024; Moreira and Pérez 2024).[14] This paper is similar in using this period of transition, but with a focus across all departments in the bureaucracy to provide empirical evidence on the personnel effects of reform. This permits me to consider how reform affects the entire cross section of the public sector, rather than quantifying the output changes within one organisation. My findings suggest that existing research that focuses on more tangible frontline service organisations for longitudinal or experimental research, regardless of internal validity, might overstate external validity with regard to top tier bureaucratic organisations. I also concur with the results from Moreira and Pérez (2024) that 'grey areas' in legislative design become the target of resistance efforts, underlining the importance of reform design to its success.

Lastly, the paper provides a novel contribution to British economic and political history. Bureaucracy, and its radical change, has traditionally been of interest to political, rather than economic, historians in the UK (Roach 1971; MacDonagh 1958). Previous scholarship has not addressed how the reform worked in practice from a quantitative perspective, instead discussing more philosophical questions on the reforms position within greater changes in British politics during the 19th century (Parris 1960). This is contrary to historical research on the American bureaucracy, which has a long quantitative tradition (Johnson and Libecap 1994a; Johnson and Libecap 1994b).[15] My results confirm qualitative accounts that the Order in Council brought about revolutionary change in the selection procedure for entry level civil servants. However, I do find that the most entrenched organs of the state resisted change most effectively, which led to a need

[11] These dynamics should also be of considerable interest to economic historians, who often refer to the importance of institutions in the rise of the West yet do so without any granular knowledge on how these institutions develop (North and Weingast 1989; Acemoglu, Johnson, et al. 2005). This would also be true of institutional (or state capacity) theory more generally in economics (Besley and Persson 2011).

[12] This also allows a more policy relevant analysis. Merit hiring in the civil service is likely important, but without an appreciation of potential pitfalls in the process policy recommendations are akin to giving a destination with no map.

[13] It is worth acknowledging the small set of research that highlights informational asymmetries as a reason why patronage may actually help public service (Voth and Xu 2019).

[14] The article by Moreira and Pérez (2024) shows that patronage continued after the Pendleton Act due to the presence of loopholes. It does however confirm that merit appointees were far more qualified and remained in the job for longer, once again suggesting these were superior bureaucrats than those appointed under patronage.

[15] There is recent scholarship that has moved towards a more quantitative approach for the UK (Xu 2018; Xu 2023), but there remains a shallow literature on the British transition to merit in the way the American transition has been focused upon (Ujhelyi 2014b; Bostashvili and Ujhelyi 2019; Moreira and Pérez 2022; Ornaghi 2016).

for follow on legislation to amend its failures.[16]

2 Historical background

In this section I briefly explain the setting of this study, the British Civil Service, as well as the reform that is specifically studied, the Order in Council of 1870.

2.1 The run-up to reform

The British Civil Service had long maintained a policy of discretionary, or patronage, appointments. Its method of appointment was entirely contingent on the exchange of social capital in return for position, junior and senior, in the pre-1855 period (Aylmer 1980; Plumb and Trevelyan 1955; Chester 1981). This led to a Victorian reform movement that sought to curtail this discretionary system as a follow up to the tremendous progress made in the electoral system (Anderson 1965; Anderson 1967; Stansky 1973; Phillips and Wetherell 1995; Aidt and Franck 2015). Proponents of a merit system championed the seminal Northcote-Trevelyan report of 1854, which concluded:

"Admission into the Civil Service is indeed eagerly sought after, but it is for the unambitious, and the indolent or incapable, that it is chiefly desired... the comparative lightness of the work, and the certainty of provision in case of retirement owing to bodily incapacity, furnish strong inducements to the parents and friends of sickly youths to obtain for them employment in the service of the Government" ("The Northcote-Trevelyan Report" 1854)

Reforms instituted in 1855 brought examinations, but only in a limited form where the opportunity to sit those exams was limited to those able to get a nomination from a given department, which perpetuated a social capital based appointment system (MacDonagh 1958; Hughes 1942).[17] Public clamour for reform, along with repeated embarrassing anecdotes such as "Hayter's halfwits" or public failures like the 1857 Indian mutiny and the review of the Crimean war, made it clear that this intermediary step had not worked.[18] In a 1869 parliamentary debate on the principle of open competition in the civil service, Colonel William Sykes recalled that while reviewing appointments as part of a select committee 'some scandalous abuses of patronage were brought before them... in one case it was stated that a lunatic had been appointed to a clerkship, and in another a greengrocer' (Parliament of the United Kingdom 1869). After the 1868 election of the

[16] In this way, the findings support the general theory 'of institutional change in the civil service [as] one of incremental alteration' proposed for the American experience (Johnson and Libecap 1994a, p.9).

[17] The exception to this is the Indian Civil Service which introduced full open access examinations in 1855 (Cornell and Svensson 2023). There is a large historical literature on this development and its effect (Compton 1968; Moore 1964; Dewey 1973)

[18] Hayter's halfwits were two men supposedly employed by Treasury Patronage Secretary William Hayter to contest for available positions with his preferred candidates, in order to give the pretence of a competition but ensuring that the preferred candidate would *always* get the position. The validity of this story is unclear, but the presence of multiple such tales about institutional negligence created a clear drag on public views of the bureaucracy. A contemporary evaluation of this period confirms this, 'we allowed party managers and public officials to decide, during the period from 1855 to 1870, who should be examined for admission to the public service. The records show that inferior and unworthy persons generally dominated (Eaton 1880, p.233)

Liberal party under Gladstone, the political climate shifted towards open access examinations, which led to the 1870 Order in Council.[19] One cannot detach this from general clamour for political reform and accountability in this period, as both the Reform Act of 1867 and Ballot Act of 1872 occurred on either side of the 1870 reform. An important point to note at this point is that reform was not popular within the bureaucracy itself (bar the Civil Service Commission which spearheaded reform efforts). Patronage brought rents and control of *who* entered a department. For the elite positions within the service, which are focused on empirically in this paper, this combination was difficult to surrender. Therefore, reform was instigated by the executive to corral them, not submitted to voluntarily.

2.2 1870 Order in Council

The reform dictated that:

> *Except as herein-after is excepted, all appointments which it may be necessary to make, after the 31st day of August next, to any of the situations included or to be included in Schedule A to this order annexed, shall be made by means of competitive examinations* (Clause V, HMG 1870)

The organisations listed in schedule A are shown in Table A1 and include the Post Office, Treasury and Inland Revenue among others. The reform was to be applied for all clerk-level positions in reformed departments. There was one carve-out given to reformed departments, Clause VII, which allowed for departmental heads to bypass the Commission if a candidates possessed 'knowledge and ability... not ordinarily to be acquired in the Civil Service' (Clause VII, HMG 1870). On its face, it would be intuitive to assume that reform was highly successful given the strong language employed. This is perhaps why historians have highlighted it as where 'patronage ceased to be of importance and the modern civil service developed (Hanham 1960, p.75). Such a strong structural break perhaps understates the importance of constant revisions over time, such as the Playfair or Ridley Royal Commissions, but even so it is widely accepted that the order led to a 'system of open competition for entry... [which] provided better-educated and more efficient public servants than patronage had done' (Roach 1971, p.228).[20] An observer who felt the move away from social connections had gone to far noted that 'if, in fact, I were asked what, in my opinion, was an essential article of the Victorian faith, I should say it was "I believe in Examination"' (Kellett 1936, p.276). However, previous research has not considered whether the policy was successful in altering selection processes.[21]

[19] Contemporary accounts suggest that there was almost universal belief among politicians that merit hiring had to be introduced to a degree. A major sticking point appeared to be that conservatives favoured a plan that would establish "twin tracks" of patronage and merit hiring, fearing that the outright imposition of merit would change 'the substance' of those in politics. One such attack was by Lord Monteagle, who argued 'that open competition was a Chinese system; and as China was not an enlightened country, the system was therefore poor' (Brown 1879).

[20] In fact scholarship of the reforms typically focused on the issues that open examinations created given their perceived dominance of the selection procedure. In particular the concern was that exams led to 'cramming' rather than broad educations which provided the soft and hard skills needed once in position. Roach points this out eloquently as the question of whether 'to train a boy to pass a highly competitive examination set on a narrow syllabus is the same thing as giving that boy a good education' (Roach 1971, p.228)

[21] Hanham, who argued for the success of the Order in ending patronage, noted himself with surprise that very little attention has been paid to the final stages of reform following the introduction of open competition in 1870' (Hanham 1960, p.75).

3 Data

Data is sourced from the *Reports of the Civil Service Commission* for the years 1864-1875.[22] These annual reports give the name, position, department, route of entry and often grades (for examined applicants) of each successful entrant. The route of entries are open competition, limited competition and no competition. The latter two are reliant on social capital, although the limited competitions still required the candidate to come first in the examination.[23] Unfortunately there are no annual records on the entire of the serving Civil Service, such as the registers in the United States (Aneja and Xu 2024; Moreira and Pérez 2024), which mean that it is only possible to observe entrants at entry rather than throughout their career. As such, the analysis is purely centred around the effects on the entry-level positions over time.

One concern with the data may be that there is systematic skew in the reporting of entry routes, creating a selection bias. For example, if law is passed that makes patronage appointments illegal then reformed departments might simply stop reporting them, rather than stopping hiring them. This would produced a biased estimate of reform effectiveness. A quirk of the civil service in the period, and the positions my sample is built from help restrict this concern. Any entrant to the Civil Service from 1855 had to register with, and be approved by, the Civil Service Commission (regardless of how you were appointed; even if your appointment was *technically* illegal) in order to gain a pension after 1859 Superannuation (Pensions) Act.[24] If a candidate was approved, as almost all were, they would be entered in the Commission's reports for that year and are thus observed in the sample. For low tier entrants the pension may have be a smaller motivation but for higher tier entrants losing access to the pension would be a far greater problem. These were individuals already making substantial sacrifices in pecuniary terms given the available returns to their skill in the private sector. In fact, the Northcote-Trevelyan report directly mentions the pension as such a draw it led connected individuals to use social capital to get their children discretionary positions in the service ("The Northcote-Trevelyan Report" 1854). The *Reports* only give clerk-level (higher tier) positions, so it is unlikely that missing bureaucrats bias the sample. On top of this, it is clear that there was no punishment for their registration, further reducing this likelihood.

I construct a repeated cross-section of all clerk-level entrants to the service, complete with an indicator denoting their route of entry.[25] This gives me a total sample of just under 10,000 individual entrants in the period 1864-1875. For the majority of the empirical analysis I keep this data at the individual level, but I am also able to collapse it at the department-year level to construct an unbalanced panel data set on departmental hiring at the clerk level.

Finally, I construct an indicator variable at the department level to proxy for status. This

[22] Unfortunately I am unable to collect data beyond 1875 because of an alteration to the examination process that pooled all examined applicants into "high" or "low" division clerks rather than giving their allocated department. As a result it is impossible to exploit within-department variation, and so I do not use the data.

[23] That being said the limited competitions were easily fixed if inept candidates were nominated alongside the preferred.

[24] They had the ability to reject certain candidates that were obviously under-qualified but at the clerk level this rejection rate was low as the main criteria was illiteracy which was a low barrier to entry (Roach 1971). This is an initial suggestion that perhaps the words of Clause V were stronger than the practical reality.

[25] To briefly mention what clerk-level means, this distinction is given by the reports which only report clerk-level entrants. One can general take the term to mean high tier, in salary and barriers to entry, positions, although there is still variation in prestige within the sample.

allows me to exploit heterogeneity amongst departments empirically. I describe this at length later in the article, but in short I construct a binary variable for membership in two groups that approximate status along historical and geographical lines. The historical group designates the most powerful departments, the 'Ancient', while the geographical group focuses on those headquartered in 'Whitehall', the seat of the UK government. It is important that these are defined by processes that existed long before reform, as one concern might be a tautological definition of power at the department level that is given by the departments capacity to escape the 1870 reform.

4 A conceptual framework

In this section I develop a theoretical framework for how meritocratic reform is implemented, how bureaucrats respond and the personnel consequences. The framework is designed to describe the transition between two states, patronage and meritocracy, and to focus upon the decision by a given managerial bureaucrat to exercise patronage for a given position as a function of derived utility.[26] I set up the environment before using the framework to derive two hypotheses on how reform that is uniform across organisations might have heterogenous implications for managers in charge of a given selection decision.

4.1 Environment

Set up. Reform is instigated upon a given organisation, nominally 'forcing' them to shift their selection procedure. Here I introduce a simple model to note how this reform interacts with strategic bureaucrats to form a new equilibrium. There is one agent in charge of selection for a position, a managerial bureaucrat, and two perfectly substitutable hiring decisions, merit or patronage hiring, for any given entry-level position. For a given position, the candidate pool comprises of individuals who would be merit or patronage hires, $x = \{m, p\}$. The principal, the government, hope that the hiring decision is carried out to maximise productivity in the organisation. I assume that productivity from merit candidates is higher, leading to the term $\tau_j(\Delta\epsilon_{j,i})$ which reflects the gain in efficiency for a merit selection – the assumption being that $\Delta\epsilon_{j,i} > 0 \quad \text{for} \quad U_{j,i,merit} \quad \text{as} \quad (\epsilon_{merit} - \epsilon_{patronage}) > 0$.[27] Managerial bureaucrats are standard utility maximisers, with managerial bureaucrat j in change of a given selection decision i maximising by

$$U_{bureaucrat} = \tau(\Delta\epsilon) + (1 - c_\theta)\alpha(p) = y + d_\theta\alpha(p) \tag{1}$$

where $\Delta\epsilon$ becomes the constant efficiency gain (loss) from merit (patronage) y. There is also a cost function incorporated into the model ($c_\theta \in [1, 0]$), which captures the potential punishments that a given bureaucrat will face if caught using patronage. This term term weights the gain ($\alpha(p)$) for that

[26] This framework therefore makes the explicit choice to set the decision with the individual manager, as in Mocanu (2023). This simplifies things and seems truer to the actual selection process, but implementing the decision at either the organisation level or combined organisation-manager decision does not affect the hypothesis derived given that the crucial component (heterogeneous power at the organisational level) is unchanged.

[27] And conversely this gain is reversed and subtracted from the utility function when a candidate is hired through patronage: $\Delta\epsilon_{j,i} < 0 \quad \text{for} \quad U_{j,i,patronage} \quad \text{as} \quad (\epsilon_{patronage} - \epsilon_{merit}) < 0$.

bureaucrat hiring a patronage candidate which I take to be a constant in the model.[28] The function is built from several components, which I consider later.

Pre-reform equilibrium. If we assume that prior to reform there are no enforced costs to patronage we have the utility function:

$$\lim_{c_\theta \to 0} [U_{pre} = y + d_\theta \alpha(p)] = y + \alpha(p) \tag{2}$$

As such, bureaucrats maximise utility with a corner solution with complete patronage so long as $\alpha(p) > y$. In other words, so long as the personal (narrowly-derived) gain is greater than the efficiency loss for a given position, patronage will be used, which in the pre-period I assume is (almost) always the case.[29] This equilibrium leads to a pervasive usage of patronage, and the necessity to exchange a form of capital for entry to the civil service.

Post-reform multiple equilibria. While previous to reform $c_\theta \simeq 0$, legislation introduces new costs through potential punishments in reformed organisations. With $c_\theta > 0$, the gains from patronage shrink accordingly as $d_\theta \alpha(p) < \alpha(p)$. The decision rule becomes whether the costs reach a critical level of c whereupon the corner solution shifts given that $U_{merit} > U_{patronage}$.[30] As such it is possible for reform to pivot the selection equilibrium if it diminishes the gains from patronage to the extent they fall below the efficiency benefits of merit hiring.[31] The solution will always be a corner, but *which* is dependent on this critical value being exceeded.

Theoretical difference-in-differences based on reform status. A managerial bureaucrat will choose to hire through merit when $\Delta U \leq 0$ which implies $d_\theta \alpha(p) < 2y$, or that the efficiency gains are greater than the gain from patronage. Of these terms d_θ is the only one that can vary due to reform as cost is imposed on bureaucrat j at the organisational level. Therefore, it would be expected that reformed organisations exhibit differential behaviour upon the passage of reform, conditional on the strength of the cost function leading to the inequality presented above, specifically they lower the share of entrants through patronage and increase merit-based hires.

[28] This decision is mainly made for ease of use – it is possible to relax it and have a function that scales with the personal gain or attitudes. The implications of this are not particularly compelling, it simply adjusts the critical threshold at which a given manager will switch from patronage to merit.

[29] This assumption is borne through by the pre-treatment averages for patronage hiring across all departments (organisations) in the sample used empirically here, shown in Table 1 and discussed later at length. It also would be fitting given existing qualitative work on the endemic reach of patronage in 19th century Britain (Hanham 1960; Eaton 1880).

[30] This solution would also work if one relaxed the efficiency gain from merit appointees so as to make it neutral, although mechanically this would require a higher cost to flip the solution – it would be logical without a productivity penalty bureaucrats would be more likely to use patronage than with one – however it is fitting with the literature on meritocratic transitions to assume patronage perverts public service (Aneja and Xu 2024; Estrada 2019). This holds both in historical settings contemporaneous to my empirical setting and in the present day.

[31] Qualitatively, this stands well with the explanation of meritocratic transitions in historical political economy, which often centre on the tipping point where patronage becomes more costly to national politicians (through several factors) than the gains they previously enjoyed (Johnson and Libecap 1994a). Obviously, in this framework the decision maker is the bureaucrat, rather than the executive, but the same mechanism is likely to operate – cost increases drive change.

Endogenising cost. The crucial part of the model therefore is the endogenous components of the cost function, which drives any equilibrium change. I introduce a simple function that endogenises the perceived cost for a given bureaucrat j:[32]

$$c_\theta = f(z) \cdot \pi_o \quad \text{where} \quad f(z) = x \cdot \delta \tag{3}$$

$f(z)$ is a linear punishment function that has a constant punishment x which is scaled by δ, which is a binary for the presence of legislative constraints. Before reform $\delta = 0$ so the entire cost function must also be equal to zero.[33] After reform, this punishment function is identical in all organisations that are 'reformed' ($\delta = 1$).

De facto power. This function is scaled by the term π_o, which is an inverse scale for the de facto power of the organisation a given managerial bureaucrat works in (the lesser status the closer to 1 an organisation o is).[34] While reform treats these organisation as uniform, instead π_o is distributed across $\pi_o \in [0,1]$ as some organisations have more capacity and status than others. The precise definition of de facto powers are those that exist *in practice* rather than in law. In a setting such as this, this commonly may relate to intra-organisational bargaining capacity that an organisation possesses. This changes the cost of using patronage if the bureaucrat knows his organisation has insulated selection from public (or politician) oversight, or is strong enough to weather any criticism. For powerful organisations this means c_θ becomes far smaller, and it is increasingly less likely that for a given managerial bureaucrat the merit corner solution ($d_\theta \alpha(p) < 2y$) is reached. Scaling that declined probability to the universe of organisations in a civil service and thousands of hiring decisions, one would expect smaller differences relative to the non-reformed organisations as π_o decreases (de facto power increases).

Institutional conditions. An important extension to this model is the degree of de jure constraint faced by organisations. I consider constraints as homogenous across organisations, there is no potential gradient within them. These can be added to the cost function as such:

$$c_\theta = f(z) \cdot \pi_d + k_i \quad \text{where} \quad k_i = m \tag{4}$$

k_i stands for the pan-organisation (institutional) strength of constraints, which takes the value of a specific positive constant m. In essence it's homogeneity provides a constant addition to the cost function, making patronage less likely. But a further consequence of this is that k_i can be great enough that it strictly dominates the rest of the function, leading to complete adherence for all (it does not vary at the organisational level). However, in mixed situations where $k_i \not\geq f(z) \cdot \pi_d$

[32] It should be noted that there are likely other variables that could be included in this function, but I chose to focus on the two given as key determinants. In the empirical setting, I consider alternative mechanisms that might explain my findings but find no conclusive evidence among these.

[33] Given that the term d_θ is $(1 - c_\theta)$ it makes sense to consider x as 1, but this could change to any arbitrary number so long as $c_\theta \in [0,1]$ or the scale is changed. Equally, the use of a 1/0 binary can also be relaxed if one wishes to introduce variation in pre-treatment function.

[34] An alternative representation of this would be to scale the punishment at the organisation level, which would lead to a similar end model, but it is more realistic to assume the same punishment across organisations, but vary its threat level by organisation, as is done here.

for only a number of organisations homogenous constraints breed unequal response.[35] Powerful organisations, despite the same constraint, evade. Weak organisations coalesce. This illustrates that in (relatively) weak institutional climates where de jure power is superceded by de facto powers at the organisational level, heterogeneity in these powers cements a negative status gradient in reform success. Persistent privileges occur because only a select group are incentivised (and able) to invest to protect those very privileges (Acemoglu and Robinson 2008). It also makes clear that strong institutional climates have far lesser concerns with this type of status gradient.

4.2 Hypotheses

The framework provides a simplified way to consider how institutional reform works across a cross-section of public sector organisations (departments in the empirical setting). It contains two major predictions on how selection changes after reform.

H 1 *Reformed organisations face an increased cost (c_θ) to selecting through patronage. Post-reform, they hire significantly more through merit than unreformed organisations.*

H 2 *The cost varies at the organisational level, so there is a distribution of c_θ that in a weak institutional climate ($k_i \not\geq f(z)\cdot\pi_o$) leads to non-compliance that is increasing in π_o. Powerful organisations use patronage as previously, while weaker organisations coalesce.*

I use individual level data from the British Civil Service to exploit a reform that allows me to observe both **H1** and **H2** using difference-in-differences for my identification. Evidence that supports both hypotheses would indicate that there is an issue with assuming homogenous treatment effects *across* the bureaucracy for such reforms in spite of general success.

5 Identification

In this section I briefly discuss the difference-in-differences strategy used in this paper. I also tackle potential issues with the core assumptions that this strategy requires.

5.1 Difference-in-Differences

To determine how policy implementation changed the structure of bureaucratic entrants, I use a canonical difference-in-differences model.[36] As such, it is possible to identify an average treatment

[35] k_i is not the term that varies here, it is the organisational powers. Thus the interpretation is that sufficiently weak institutional conditions (that are homogenous) lead to a position where some organisations possess the capacity to circumvent reform efforts.

[36] There has been significant tumult in the literature on this research design, particularly towards the inaccurate estimates that TWFE can provide in certain empirical settings (Chaisemartin and D'Haultfuille 2020; Goodman-Bacon 2021). For the chosen policy, this is not an issue as all units are *treated* at the same time and there are no switchers out of treatment.

effect on the treated (ATT) for treated departments using the following linear probability model:

$$\text{patronage}_{i,t,d} = \beta_1 \times \text{reformed}_d \times \text{post}_{i,t} + \tau_t + \theta_d + \varepsilon_i \tag{5}$$

The coefficient of interest (β_1) is the interaction term between reformed and post. This should be interpreted as the change in probability (in percentage points) that an individual i in treated department d in a given year t enters the civil service through patronage, relative to the control group post-treatment. Reformed is a dummy variable for a department being directly reformed, while post takes the value 1 if the year is 1870 or later (1870 being the year of treatment). It is possible that estimates might still be biased in a 'naive' policy analysis without fixed effects due to unobserved variation at the departmental or hiring cycle level. To provide a causal estimate I therefore use departmental (θ) and year (τ) fixed effects. Intuitively, these allow the analysis to focus on *within-department* variation, an important step given that (for example) the Post Office may exhibit substantially different hiring practices to the Irish County Surveyor's Department. Equally, there may be large shocks to demand in certain years (hiring freezes, census-year surges) that lead to cross-bureaucracy behaviour adjustments. I always cluster at the department level, as this is the unit at which treatment occurs (Abadie et al. 2022).

5.2 Assumptions

For a causal interpretation of the difference-in-differences estimates, the natural experiment must satisfy certain conditions. Spillovers are less of a concern here, as they are positive – i.e. the introduction of reform and a standardised open examination process leads untreated departments to provide some open access roles. There is no possibility of a negative spillover in an environment where everyone is pre-treatment appointed through patronage, which can be seen graphically in Figure A3. As such it is almost certain all estimates of the ATT will in fact be *downwards biased* of the true average treatment effect (ATE). The main concern therefore for a causal interpretation is selection, in particular selection into treatment. Departments might select into, or out of, treatment due to their underlying perspective on meritocratic recruitment.[37] Departmental fixed effects mean that level differences are not problematic, but selection into treatment could lead to the ATT being biased if it picks up on departments willing to change (and thus who selected into treatment). For a causal interpretation, more precisely, it must be assumed that:

Identification Assumption: *Departments may differ in their baseline hiring behaviour, but these differences are unrelated with the timing of reform:* $E[\textit{reformed}_{i,d,t} \times \textit{post}_{i,d,t} | \varepsilon_{i,d,t}, \theta_d, \tau_t] = 0$. *Systematic differences are acceptable (absorbed by the fixed effect structure), so long as they do not interact differentially with treatment.*

Which departments were reformed? To assess this assumption, it is important first to focus on which departments were reformed in 1870, given that public sector reform is typically a *deliberate*

[37] Policymakers might even deliberately target certain organisations, providing non-random treatment assignments. This is often the case for policy choice that is centred around deliberate thresholds, such as the Pendleton Act (Aneja and Xu 2024).

rather than *idiosyncratic* policy choice (Ujhelyi 2014a). There was no single arbitrary rule followed, such as population thresholds rule in the Pendleton Act. As such, it requires careful examination to show whether treatment was random. In Table A5 I list departments that were directly affected by the reforms of 1870. Some historians have argued that the notable lack of 'the ancient departments of the Home and Foreign Office, which fought hard to retain patronage on the threatened basis of character' might explain treatment assignment (Davis 2006, p.30).[38] However, such a view fails to explain the multitude of non-treated departments that were not ancient departments (such as the Census Office or Metropolitan Police) nor that ancient departments such as the Admiralty, Treasury or War Office are given as treated organisations.[39] This variety in both treated and untreated groups also clouds an idea that reforms might deliberately avoid departments where important informational asymmetries might be cured through patronage networks (Roach 1971).

The idea of a quasi-random selection is supported by evidence that assignment was dictated by bargaining and chance decisions within the executive (Liberal Party) that were orthogonal to the wants of a bureaucracy unanimous in its desire to protect patronage.[40] Fry argued that William Gladstone (Prime Minister at the time) avoided resistance in his cabinet through 'a neat compromise whereby all branches of the Service were thrown open to competition where the Minister concerned agreed' (Fry 1969, p.69). That this occurred at the ministerial level is important for identification purposes; decisions were made by the executive for the bureaucracy, rather than the bureaucracy (and specific departments) having a role in deciding their own fate. Parliamentary debates at the time of reform offer some more clarity upon this. Henry Fawcett (MP for Brighton) raised a concern on why merit had not been implemented universally, with the minutes recording him as objecting to that fact 'that officers in the Foreign and Home Departments were excepted from the Order in Council' (Parliament of the United Kingdom 1870b). The response from the Chancellor of the Exchequer (Robert Lowe) makes it clear that this executive compromise was the rule followed: 'They consulted every head of a Department on the subject, taken his opinion, and acted upon it. The heads of the Home Office and of the Foreign Office thought it would be better the system should not be extended to them. To that they saw no objection, because the great mass of the Civil Service were willing to come in. The matter was in the nature of an experiment' (Parliament of the United Kingdom 1870b).[41] These debates make clear that the Liberal Party viewed this initial action as a natural experiment for the principle of meritocracy, with Lowe even stating 'the thing would be very shortly tested, and then it would be seen whether it could be safely extended' (Parliament of the United Kingdom 1870b).

[38] The ancient departments of the Civil Service were, and still are, the most storied and prestigious in the UK - I return to this distinction later in the article.

[39] The presence of the Treasury is particularly important to this counter. Over the eighteenth and nineteenth centuries the Treasury had consolidated power, as the convergence of 'parliamentary and Treasury interests [allowed it] to become something than can be meaningfully called 'control' ' (Roseveare 1973, p.73). As such its presence as a treated department suggests it would be incorrect to attribute treatment to an unobserved factor such as departmental status.

[40] The desire of the bureaucracy to protect patronage was touched upon in an American columnist's account of the 1870 reform, where he stated the importance of politicians to the process stating 'any ministry would be loath to give up such a privilege' (Brown 1879)

[41] Heads of a Department are the elected politicians, rather than the leading bureaucrat, which supports Fry's account. This was made clear by Gladstone when answering a question from Lord Claud Hamilton on the reason behind the exclusion of the Foreign and Home Office's from the reform. Gladstone refers to the Home Secretary's decision, making clear he is the Head of the Home Office (to refer to Lowe's remark), and that the opinion of this individual determined assignment, whereas the Admiralty, Colonial Office and Treasury's elected leaders took alternative positions (Parliament of the United Kingdom 1870a).

Table 1: Balance test: Difference-in-means

	Control mean (1)	Treated mean (2)	Difference (T-C) (3)
Route of entry:			
Patronage	0.989	0.998	0.00858**
	(0.00385)	(0.000793)	(0.00393)
Limited Exam	0.297	0.297	-0.000699
	(0.0169)	(0.00743)	(0.0185)
Direct Appoint.	0.692	0.701	0.00928
	(0.0171)	(0.00745)	(0.0186)
Observations	4,510	4,510	4,510

Notes: The table presents a balance test between untreated and treated departments. Using individual level data I calculate the mean for each route of entry to the Civil Service pre-treatment and then use a difference in means regression to compare the means. The route of entry is given in the rows, and for that given row the binary variable takes the value one if the appointment is by that route of entry. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

There are three cases where this idea of idiosyncrasy might fail, and selection into, or out of, treatment appears possible. While chosen by their elected leaders, the Foreign and Home Office do appear to have used the grounds of their work to actively justify exclusion from the law, whereas the other untreated departments are much quieter participants in the process. On the other side, the Civil Service Commission, as the driver of reform, clearly opted into treatment. Entrants from the three departments account for 2% of the sample from my preferred specifications, and all results hold when they are excluded from the sample, given in Table A11. From a qualitative perspective, treatment assignment appears to be driven by bargaining idiosyncrasies within the executive, thus meeting the identification assumptions. One can also look at potential differences further through some simple empirical tests on pre-treatment data.

Pre-treatment differences. In Table 1 I carry out a simple difference in means regression to test for balance between treated and untreated departments. To do this I split my pre-treatment sample by the route of entry: (1) any form of patronage, (2) examinations by nomination, (3) direct appointment. There is high overlap therefore as (2) and (3) constitute (1). For sample split I regress an indicator for reform to see whether there were any substantial pre-treatment differences between the groups. My main interest is the possible distinction between (2) and (3), as differences based on how patronage was used might reveal something about the departments themselves.

In columns 1 and 2 I give the means for the control and treated groups respectively, with the difference between the two given in 3. This difference-in-means estimate gives a crude way to compare pre-treatment trends through asserting balance between the groups.[42] I find that there is only one significant difference, the absolute number of entrants through patronage. However, the difference in means is less than 1% point given how few enter through merit in the pre-treatment

[42] I also do this in a more sophisticated way using event study specifications in order to test for pre-trends. The results, given in Table A6 and Figure 1, reject any pre-treatment effects.

period (99% enter through patronage). The direction of the relationship further suggests this is not a concern. The direction of any bias would point downwards, rather than upwards, as the results suggest reformed departments were more likely to hire through patronage before being treated. What is more important for balance is that departments used patronage in similar ways. Making a distinction between the two routes of patronage (limited examination or direct hires) in the latter two rows of the table, I find no significant differences between departments, with both treated and control groups exhibiting remarkably similar behaviour. This confirms that departments exercise patronage in the same way pre-treatment, leaving no concerns that there may be some differentials in their behaviour that affect selection into treatment or the parallel trends assumption. I also do not find evidence that the departments hired candidates of different quality, as proxied by their grades in examinations. By collecting all available examination grades for successful candidates pre-treatment, I show in Figure A2 there are no significant differences by treatment-control status pre-treatment.

6 Impact of reform on patronage

I initially focus on whether there was substantive change to the selection process in the British bureaucracy after reform. As Table 1 shows, before reform the standard method of appointment was through patronage (99% of entrants). The graph in Figure A3 shows a large climb in merit hires in the treated group, which supports the hypothesis descriptively.[43] I estimate the causal effect of reform on the probability (or share, depending on specification) of entering through patronage using Equation 5. In line with **H1**, I find that reform brought about a significant fall in the number of applicants entering treated departments through patronage.

LPM Estimates To test my first hypothesis of general reform success, I estimate a linear probability model using the exact specification in Equation 5. A negative coefficient would indicate that treated departments were less likely to use patronage than the control group post-reform.[44] The results are presented in Table 2. I report the regressions in columns 1-3 with differing fixed effects specifications, although my preferred specification is the third given the more robust fixed effect structure. The results are stable across the specifications, suggesting that the probability a given candidate entered through patronage was between 34-40 percentage points lower in treated departments post-reform relative to the control group. This is clearly a large economic effect, as it indicates a decline of over one-third from the original mean pre-treatment.

I also use an 'event study' specification to partially validate the parallel trends assumption. One concern (as an example) might be that treated departments may have been more likely to appoint through merit pre-treatment, leading to a biased effect that is unrelated to reform (although the descriptives suggest this is not the case). Another might be there is anticipation, where treated

[43] It also shows strong evidence of positive spillovers, which would suggest any difference in differences ATT is downwards biased of the true ATE. While this attenuation is not a problem for the causal interpretation, it should be noted when considering the magnitudes of the effects.

[44] Specifically, the interaction term β_1 gives the change in probability of a given route of entry appointment in the treated departments relative to the control departments post-reform.

Table 2: Baseline DiD

	Dep. Var: Patronage (1/0)		
	(1)	(2)	(3)
Reform × Post	-0.408***	-0.374***	-0.341***
	(0.125)	(0.128)	(0.120)
Observations	9,309	9,272	9,272
Dept FE		✓	✓
Year FE			✓
Mean Dep. Var.	0.788	0.787	0.787

Notes: The regression model is a linear probability model, where I regress the interaction term (reformed * post-treatment) upon a binary variable for patronage use for each individual appointment from 1864-1875 in the British civil service, thus estimating whether reformed departments changed their manner of appointment post reform. I employ fixed effects at the department level to minimise individual quirks in the hiring process.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

departments realise they will be reformed and so adjust their behaviour accordingly (in this situation hiring more through patronage in its last days). I consider the effect of treatment in all years through the sample relative to the last period pre-treatment (1869). These help to validate the parallel trends assumption if the test for joint significance of coefficients pre-treatment is not different from zero. They also show the time dynamic of any effect found. The results are presented in Table A6. They show no evidence of pre-trends, p-values are between 0.4-0.74 depending on specification, validating parallel trends as I cannot reject the null hypothesis that these periods are jointly equal to zero. The time dynamic indicates a fairly constant effect from 1871 (bar one year in 1874 that is not significant), suggesting that departments quickly complied with new legislation on average and exhibited persistently higher declines in patronage use.[45]

Panel Data Estimates I also use a departmental panel to complement these results. I collapse the individual data at the department-year level to create an unbalanced hiring panel. As such, the focus changes from individual level decisions to how departments at large shifted their selection procedures at an aggregated level. For each department-year I calculate the share of entrants through patronage, which I regress upon the same treatment indicator using the equation:

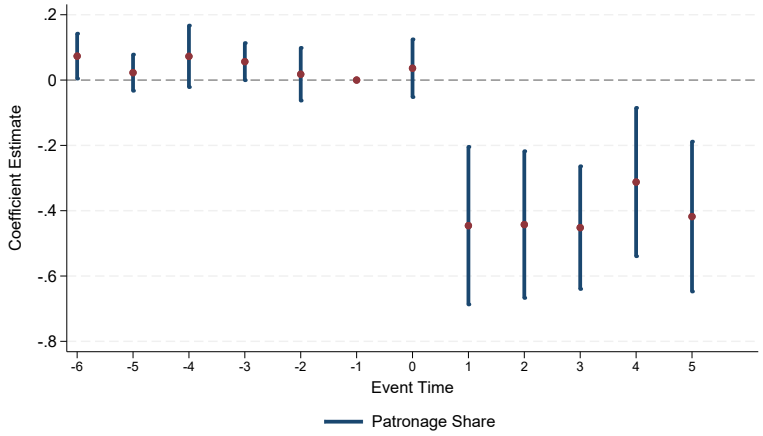
$$\text{patronage_share}_{d,t} = \beta_1 \times \text{reformed}_d \times \text{post}_{d,t} + \tau_t + \theta_d + \varepsilon_d \tag{6}$$

The results are reported in Table A7. I adopt the same varied fixed effects approach as in the baseline with similar results. This is reassuring, as differences between individual and department level regressions would be surprising. The β_1 coefficients from Table A7 are remarkably similar to those

[45]The reason for the insignificant coefficient in 1874 relates to spillovers. In this year there was a spike in merit hiring in the control group of departments, which understates how reform had changed selection in the treated group in the year. As mentioned before, one should treat all these estimates as attenuated relative to the true change engendered in treated departments if the control group had been left in isolation (the ATE).

presented previously. The effect ranges between a 33-38 percentage point fall in patronage shares among treated departments. They suggest once again that reform had an immediate effect on how departments hired, reducing their intake through patronage substantially.

Figure 1: Dynamic effects of reform on patronage share



Notes: The figure presents the coefficient on the regression in Table A7 column 3 but interacting the treatment with each individual year (omitting the final pre-treatment year, 1869, as the reference category). The dotted line is at a 0 coefficient, red dots give the coefficient and the blue bars the 95% confidence intervals of said coefficient. Treatment occurs at $t = 0$.

I also estimate event studies using this panel, which allows one to observe both the dynamic nature of the effect and potential pre-trends that would violate the identification assumption as in Table A6.[46] In Figure 1 I plot the results of this. There are some minor pre-trends, notably in period -6 (1864), which suggest that treated departments used more patronage before reform. This is not a concern. The estimate is small (but precise given the lack of non-patronage hires pre-treatment) and a test for the joint significance of all pre-treatment coefficients cannot reject the null hypothesis that they equal zero. The treatment effects are once again large, suggesting qualitatively similar drops in the share of entrants through patronage relative to the control group after reform. As in the event study LPM specification, the effect appears in 1871 the year after reform was passed, confirming immediate adaptation to new regulations that persists.

Robustness This result holds up against a variety of different robustness checks. An advantage of the panel structure is that it allows me to use the doubly robust estimator introduced by SantAnna and Zhao (2020) to relax the assumption of unconditional parallel trends by conditioning on pre-treatment selection behaviour. This allows me to refine the comparisons being made further by using inverse propensity score weighting. In Table A8 the ATT is slightly reduced, suggesting a 26 percentage point drop in patronage among treated departments which is around 8 percentage points smaller than the effect found in Table A8. The reason lies in how the estimator works, as it requires a balanced panel. This leads me to exclude smaller departments that did not hire annually. If these smaller departments were more likely to comply with legislation (as my results

[46] These regressions follow the equation: $\text{patronage_share}_{d,t} = \beta_1 \cdot \sum_{t=1864}^{1875} \text{event_time} \times \text{reformed} + \tau_t + \theta_d + \varepsilon_d$. The omitted year is 1869, the year prior to treatment.

here and prior hypothesis suggest) then the effect would shrink when they are removed from the sample.[47] My results are further robust to using nearest neighbour, conventional propensity score or coarsened exact matching – all done using the same pretreatment averages as the matching input – with the results given in Table A9. They are also unaffected by converting the baseline from a LPM model to logit/probit models (and using jackknife resampling), as shown in Table A10.[48] One final concern might be about the comparison being made, in particular that certain parts of the treatment and control groups might be 'unfair' comparisons if they were either selected into or opted out of treatment. I note earlier that there are three potential departments that act in such a capacity, the Civil Service Commission (selects into) and the Foreign and Home Office's (opts out of), and so results should hold without the inclusion of these organisation. The results can be seen in Table A11, and are slightly larger than the baseline, likely due to an attenuation introduced by improper treatment-control assignment.

7 Explaining variation in compliance

The above section makes clear that policy had clear positive aggregate effects on treated departments. However, after reform treated departments, on average, still hired around 40 percent of their entrants through patronage, despite its supposed illegality. **H2** suggests that compliance will be related to departmental-level de facto powers, which I test empirically in this section. I show that de facto powers among the most elite and powerful departments, the 'Ancient' organs of the state, allowed them to avoid changing their previous patronage use – using both loopholes and directly disobeying legislation to do so. I also look at a few other potential explanations, finding no convincing evidence in their favour.

7.1 Defining power in the British Civil Service

De facto power in an organisation means the actual, regardless of legal status, capacity (or authority) that the organisation holds. In the case of the public sector and reform, the main importance of the concept regards intra-organisational interactions. Reform compels change through legislation, but uses regulatory bodies to enforce it. In this setting, the body charged with enforcing reform was the Civil Service Commission (although they lacked certain crucial powers as mentioned before). De facto power here manifests in the capacity to insulate selection processes from Commission oversight, to directly flout legal changes, or to exploit gray areas in selection that require bargaining power. Defining organisations that hold this power is *a priori* challenging. It is important to avoid tautologies, as one may inadvertently define power through the capacity to avoid reform, creating a circular logic. To avoid this, I focus on pre-treatment measures of status that likely reflect increased intra-organisational bargaining capacity, and so may drive non-compliance. To do so, I

[47] The event study using this estimator, shown in Figure A4, indicate that this smaller estimate mainly comes from the fifth post-treatment period (1874) where there was a null effect (similar to the fifth period in Table A6). As before, this specific year is clouded by positive spillovers in the control group.

[48] For obvious reasons these types of models are not preferable to the LPM specification as they prevent use of a TWFE specification while using individual-level data. It is still worthwhile to confirm similar sign direction.

exploit the historical development of the British Civil Service to create two traditional groupings of departments: Whitehall and Ancient. Whitehall departments are those in Whitehall, which is the physical centre of the British Civil Service. Ancient departments are all in Whitehall, but they are separated by a further degree of prestige, historical importance and role in British governance. As such the distinction offers a stepped classification of pre-treatment status amongst organisations at the time. All departments that do not fit either category serve as the implicit reference group.

To operationalise these distinctions I create two binary variables. Whitehall takes the value one for departments that are located in Whitehall, and zero otherwise. Ancient takes the value one for Ancient departments, and zero otherwise. The exact departments included in each are given in Table A12. There are treated and untreated departments that take the values one and zero in both instances.

7.2 Power and patronage use

To consider this statistically, I use a triple difference specification to compare the change in patronage among treated groups based on whether they are in or out of the status groups explained above:

$$\text{patronage}_{i,t,d} = \beta_1 \times \text{reformed}_d \times \text{post}_{i,t} \times \text{group}_d + \tau_t + \theta_d + \varepsilon_i \tag{7}$$

As such the coefficient of interest, β_1, gives the percentage point difference in the probability of hiring through patronage (given that outcome) between treated departments within the group and those outside of it.[49] For example, if $\beta_1 > 0$ when $\text{group}_d = \text{Ancient}_d$ it suggests that relative to non-Ancient treated departments the Ancient treated departments hired more through patronage post-reform. My expectation is that more powerful departments will have a higher probability of patronage use after reform, meaning $\beta_1 > 0$.

The results are presented in Table 3. The coefficients on both Whitehall and Ancient are significant, economically and statistically (columns 1-4). They suggest a 43-50 percentage point increase in the probability a candidate enters the civil service through patronage in treated departments belonging to one of the groups relative to non-group treated departments after reform. In line with the prior hypothesis, the lower tier group in terms of prestige, Whitehall, exhibits smaller effects in both magnitude and significance. This suggests that the result is driven by the Ancient groups, which would be in line with the prior hypothesis. To test this formally, I separate the Ancient and Whitehall groups into exclusive groups (i.e. I remove all Ancient groups from the Whitehall grouping) and include both as triple difference terms simultaneously. In essence this horse races the two to see which *drive* the effect found. The results are given in column 5 and 6, which make clear that the Ancient departments are those which resist change more, with around a 50 percentage point higher probability of patronage after reform (for a given selection) in both specifications. In the final two columns (7 and 8), I restrict entry into the treatment group to just Ancient departments and run the baseline model (all treated non-Ancient departments are dropped from the sample). The coefficient therefore gives the change in probability a candidate enters through patronage (in percentage points) between treated Ancient departments and the entire control group of departments post-

[49] group_d being the binary variable for whether a department is in the Whitehall or Ancient grouping.

Table 3: Power and patronage use post-reform

	Dep. Var: Patronage (1/0)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
R × P × Whitehall	0.362*	0.430*						
	(0.187)	(0.224)						
R × P × Ancient			0.494***	0.507**	0.493***	0.521**		
			(0.178)	(0.199)	(0.184)	(0.211)		
R × P × Whitehall (non-Ancient)					0.102	0.210		
					(0.192)	(0.253)		
R × P (Ancient only treated)							-0.0775	-0.114
							(0.114)	(0.117)
Observations	9,272	9,272	9,272	9,272	9,272	9,272	2,725	2,725
Dept FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE		✓		✓		✓		✓
Mean Dep. Var.	0.788	0.788	0.788	0.788	0.788	0.788	0.922	0.922

Notes: The regression model is a linear probability model, where I regress the interaction term on a binary variable for whether an applicant enters through patronage or not. The interaction term is a triple interaction between my conventional interaction (given R × P for brevity) and a binary variable for the group (given in the row title). The final two columns report a regression of the conventional difference in difference term (R × P) with the treatment group just including Ancient departments. I employ fixed effects at the department level to minimise department-level quirks in the hiring process. I also use annual fixed effects for a similar reason.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

reform. This allows me to see whether Ancient departments change their behaviour relative to the unreformed, given that the previous regressions compare them to only reformed departments. The point estimate is slightly negative (minus 7-11 percentage points) but is not distinct from zero. This makes clear that not only were Ancient (powerful) departments able to avoid reform relative to other treated groups, they were able to act in a manner undistinguishable from untreated departments. As predicted in **H2**, reform had changed the masses but not the elite.

How did non-compliance occur? The results show that this status gradient in reform success (compliance), but it is also interesting to consider how elite departments avoided legal constraints.[50] My setting allows me to disaggregate the types of patronage used by departments, a rare opportunity for empirical research. After reform there were three types of patronage still available to departments: (1) direct, non-examined, appointments, (2) examinations that required nomination, (3) direct appointees under Clause VII protections.

I expect that (2) is unlikely to be used post-reform, as limited examinations were directly supervised by the Civil Service Commission and patronage use through this system would draw considerable attention. (1) and (3) are both 'grey area' routes that would be more attractive to exploit post-reform. (3) is clearly defined, a department had to apply for an exception which could be granted by the Civil Service Commission based on the claim that they needed patronage to hire the *right* candidate. This would be clearly manipulable, and so I would expect this to vary with status at the organisational level because that would likely permit greater inter-organisational bargaining power.[51] (1) is less clearly defined, as it regarded appointments made without applying for this Clause, making it more directly disobedient. The reason given for the use of this in contemporary accounts is the existence of historical Acts of Parliament granted to several departments in the Civil Service. The Acts of Parliament gave selection rights to the department, and were a higher form of legislation than the 1870 reform. This hierarchy of legislation meant that for several departments (a exhaustive list of which does not exist) it was possible to invoke legal precedent to maintain typical patronage appointments to do so.[52] However, such an action brought with it clear scrutiny and potential repercussions from the government. If these acts were more likely to be granted to higher status (older) departments, and these departments were also more able to withstand the political costs of their use, one would also expect (1) to be more prevalent along the status gradient. I expect that (1) is more challenging to exploit than (3) given that it required substantial political capital to both possess an Act of Parliament and to defend its usage amongst widespread political opposition. As a result, the only distinction between Ancient and Whitehall should be that Ancient departments are more likely to use (1), while both use (3) equally.

I use the same triple difference specification as before, but change the outcome variable to the disaggregated measures of patronage. Thus, my coefficient of interest here gives the percentage point difference in the probability of hiring through the given measure between treated departments within the group and treated departments outside of it. My expectation is that Whitehall and Ancient departments exploit loopholes under Clause VII (3) but only Ancient departments exploit the grey area of disobedience (1). I do not expect there to be a difference based on limited competitions (2), which were directly controlled by the selection body.

[50] Existing research on the Pendleton Act finds that loopholes (or gray areas) in reform could be exploited to maintain patronage privileges (Moreira and Pérez 2024). Other research that focuses on the New York Police Department suggests that more blatant disobedience was possible (Leucht 2023).

[51] This concern was noted in debates over the Order in Council by Scottish MP Roger Sinclair Aytoun, who argued that Clause VII introduced 'another species of patronage of a not less prejudicial character... namely, the appointments which might be made under the 7th section of the Order in Council that he had quoted, and which the Government might confer on their friends.' (Parliament of the United Kingdom 1870b).

[52] These acts are mentioned in a series of Parliamentary questions asked by Major Myles O'Reilly in 1873. The Chancellor responded to his question on the ability for certain departments to avoid merit hiring by stating 'the Treasury and all other Departments to whom full powers of appointment were granted by Act of Parliament possessed the right [to do so]', alluding to an question from the week before where he acknowledged that 'no Order in Council could alter or repeal, or in any way vary an Act of Parliament' (Parliament of the United Kingdom 1873a; Parliament of the United Kingdom 1873b).

Table 4: The continued use of patronage

Dep. Var:	Limit. Comp (1/0)		No Comp (1/0)		Clause VII (1/0)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
R × P × Whitehall	0.0852		0.351		0.332***		
	(0.120)		(0.258)		(0.0745)		
R × P × Ancient		0.0369		0.477**		0.282***	0.313***
		(0.129)		(0.235)		(0.0675)	(0.0733)
R × P × Whitehall (non-Ancient)							0.406***
							(0.120)
Observations	9,272	9,272	9,272	9,272	9,272	9,272	9,272
Dept FE	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.176	0.176	0.612	0.612	0.0304	0.0304	0.0304

Notes: The regression model is a linear probability model, where I regress the interaction term on a binary variable for whether an applicant enters through patronage or not. The interaction term is a triple interaction between my conventional interaction (given R × P for brevity) and a binary variable for the group (given in the row title). The final two columns report a regression of the conventional difference in difference term (R × P) with the treatment group just including Ancient departments. I employ fixed effects at the department level to minimise department-level quirks in the hiring process. I also use annual fixed effects for a similar reason.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

I present the estimates in Table 4. In line with my expectations, I find that neither group has a larger probability of using patronage through limited competitions in columns 1 and 2, where the estimate is effectively zero. In columns 3 and 4 however, I do find a substantial relationship with direct patronage use. Both groups report economically large point estimates (35-47 percentage point increased probability), but only the Ancient result is distinguishable from zero (column 4). This is in line with my prior expectations, only the most powerful can exploit this particularly challenging grey area. I also consider the usage of Clause VII, a loophole that allowed departments to directly appoint a candidate if they invoked that no one of equal quality could be found. In columns 5 and 6 I find initially that both groups have large (and significant from zero) point estimates. In column 7, I horse race the two as before, comparing the effects for the Ancient and non-Ancient groups in the Whitehall grouping. Both coefficients grow slightly and remain significant, suggesting that both tiers of prestigious departments held the bargaining capacity to invoke loopholes and exploit gray areas in reform design.

Taken as one, these results show how status amongst allowed patronage to persist after reform. Ancient departments were able to circumvent new legislation through direct disobedience and strategic use of a given loophole that was possible given their bargaining capacity. At a lower level of power, Whitehall departments could also use bargaining to their advantage relative to the rest of the treated departments regarding the loophole, but lacked the outright power to ignore. Both groups avoided using patronage in more tightly supervised environments. The results make clear that the complicated layering of authority amongst departments produced grey areas, which are often present in any reform effort. These grey areas are best exploited by departments that hold status, and thus an advantage in intra-organisational bargaining.

Qualitative support. Qualitative evidence, from the period and historians of the topic, further corroborates this mechanism. As Fry notes, it was only by 1894 (24 years after the initial legislation) that 'the Treasury felt able to say that open competition for the Higher Division was 'now the rule' ' (Fry 1969, p.69).[53] The imperfect nature of reform led Richards to note that the post-1870 equilibrium was one where 'patronage had to be justified and perhaps excused; it could not be assumed' (Richards 1963, p.54). Alongside prior qualitative records mentioned in this article, it appears clear that there was a contemporary consensus that reform had not worked perfectly. My analysis helps to show that the main issue, as several contemporary politicians noted, was the ability for more powerful departments to exploit loopholes and disobey without punishment.

Another way to consider this is how legislation responded to the 1870 attempt. In 1875, the Ridley Commission filed a report on the success of the 1870 reform – with contemporary authors noting the need for 'the abolition of the system of having different grades of offices... the clerk of the Foreign Office to be no longer able to sneer at his fellow-clerk in the Post Office or Customs' (Bowie 1875, p.284). The commission's main scepticism with the system of open competition introduced in 1870 was clear with the introduction of a homogenised selection process where clerks were centrally appointed through two schemes from which successful candidates were allocated to departments (White 1932). The success of this additional measure cannot be tested given the available data, but its existence suggests that contemporaries quickly became aware of the capacity for elite departments to resist reforms – equals under law but first among those in practice.

7.3 Alternative explanations

It is possible that there might be some confounding mechanism that influences the results to suggest that the most prestigious departments are least effected by reform. If this were to be true, **H2** is not correct and there is instead an alternative latent variable that influences reform response. Here, I briefly tackle two potential mechanisms that could influence the results presented previously. I find no substantive alternative to my hypothesis.

Supervision costs. One might be concerned with supervision costs as a potential confounder. This would work if it were possible that departments with higher de facto powers were also more costly to supervise then results might be biased. If one was to argue that supervision costs come from the power of a department, then this speaks to the mechanism that I lay out before. If one was to argue that these costs come from proximity, so that departments further from the nexus of reform are less easily constrained, then this would also not be of relevance. All Ancient departments were based in Whitehall, as was the Civil Service Commission. Therefore costs cannot have been higher due to proximity unless costs were a function of departmental power (which would speak to my proposed mechanism). To entertain the possibility that supervision costs through proximity might provide a complementary mechanism alongside my hypothesis, I estimate a triple difference model that separates the treatment effect based on whether a departments work was focused in the UK or abroad (domestic or foreign). I show the results in Table A13, which suggest that there are

[53] In fact, Fry disputes this statement as not being entirely accurate in 1894 after several modifications to the original legislation, let alone in 1870.

no significant differences in the treatment effect based on general supervision costs for regulators.

Positional prestige. There also may be variation in the sample in the types of position offered. For example, it may be true that more prestigious departments offer on average higher prestige positions. If it were instead true that positional prestige is crucial to perceived cost (rather than departmental power as claimed) then controlling for the type of position should lead to a diminished point estimate.[54] I add a control variable into my LPM specification that directly controls for any variation related to positional prestige among departments. I do this through two different methods, directly controlling for all positions individually (which is problematic both due to changes of name for the same positions between departments and over time) and by respecifying a position variable that takes the values 1-3 for clerks (strict), clerks (weak) and others.[55] I control for these variables in the baseline regression from Table 2 to hold fixed this potential source of bias. The results are shown in Table A14, and shrink slightly but remain significant at conventional thresholds, suggesting that potential discrepancies in positional prestige cannot explain the treatment effect found. This rules out the possibility that elite departments do not change because they offer more prestigious positions.

8 Conclusion

In this paper, I focus on how meritocratic reforms affect the *entire* cross-section of the public sector. I propose a novel theory that the status gradient between organisations in the public sector leads to de facto power inequalities that affect the success of reform. To consider this empirically, I study a unique public sector reform in the UK where organisations were treated for idiosyncratic reasons across the *entire* bureaucracy to study where reform achieved success and where it failed to compel organisations to change. Rare data at the individual-level allows me to compile data on the universe of entrants to high-tier positions and explore this heterogeneity without sample constraints. I initially estimate clear policy success, finding that reformed departments used patronage far less after reform. This result is robust to several respecifications, confirming my first hypothesis that reform leads to significant increases in merit hiring among reformed departments. However, in line with my second hypothesis I find a clear status gradient in reform success. Patronage use was far higher in treated prestigious departments post reform, and showed no significant difference to the control group. I find that this was driven by the most elite (Ancient) departments in a triple difference specification. Further tests show that this occurred due to the capacity for elite departments to exploit grey areas, and rely on organisational bargaining capacity to hold off any repercussions. While legislation treated them as equals, those with power were in fact first among these. This directly supports my theory, and shows that ambitious reforms in the public sector can

[54] To imagine this more specifically, take the previous model given in this paper and consider instead that what scales $f(z)$ is not π_o but instead the variable ρ_c, the level of prestige for the position being selected for. In this case, managerial bureaucrats are incentivised to use patronage more when ρ_c is higher as these are applicants who are more likely to take senior positions in their organisation so control over them is both personally satisfying and allows them to derive higher rents.

[55] Clerk strict is for those who title was just “clerk" while clerk weak is those who included other words, such as “regional clerk" or “railway clerk". Others is for neither group.

fall into a trap of taking the lower hanging fruit.

The results provide unique evidence to a literature across political science and economics that has to this point been more focused on the consequences of patronage or its reform. This literature has considered several externalities, and even why reform itself is instigated. However, it has not yet considered implementation across organisations in a general equilibrium sense. My article provides an empirical setting where this is feasible, and shows that general equilibrium considerations are important, as partial equilibrium results in frontline service organisations may not hold external validity to more insulated, higher status organisations (that are also more difficult to observe empirically from a personnel or output perspective).

This provides a generally discouraging example for policymakers in weaker institutional climates. Existing research on frontline organisations, regardless of its internal validity, should not simply be extrapolated across a bureaucracy given the heterogeneity in organisations contained within.[56] From a cynical perspective, it might suggest that such a climate predetermines the limits for meritocratic reforms. In politically stable climates, such a predicament is not necessarily a problem in a long-run perspective. As in the UK, successive reforms (done over a moderate time period) that directly correct for previous failings can eventually provide a solution. However, in climates where one step forward may be followed by two backwards given instability there is a greater importance for immediate policy success. Optimistically, the British case offers the recommendation that policy must be implemented under an supervisory body with power to both explore and enforce potential non-compliance. This may not be enough to prevent the emergence of a status gradient in adherence but it may shift the internal calculus of patronage use for enough bureaucrats to make a substantive difference.

[56] One should note here that this is only when considering extrapolation to the higher tier organisations. Meritocratic reform has been shown to deliver large positive returns in several contexts outside of such organisations, and it is not the intention of this paper to call into question the efficacy of this outside of a narrow group of elite organisations.

References

ABADIE, A., ATHEY, S., IMBENS, G. W., and WOOLDRIDGE, J. M. (2022). "When Should You Adjust Standard Errors for Clustering?*". *The Quarterly Journal of Economics* 138.1, pp. 1–35.

ACEMOGLU, D., JOHNSON, S., and ROBINSON, J. (June 2005). "The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth". *American Economic Review* 95.3, pp. 546–579.

ACEMOGLU, D. and ROBINSON, J. (Mar. 2008). "Persistence of Power, Elites, and Institutions". *American Economic Review* 98.1, pp. 267–93.

AIDT, T. S. and FRANCK, R. (2015). "DEMOCRATIZATION UNDER THE THREAT OF REVOLUTION: EVIDENCE FROM THE GREAT REFORM ACT OF 1832". *Econometrica* 83.2, pp. 505–547.

ANDERSON, O. (1965). "The Janus Face of Mid-Nineteenth-Century English Radicalism: The Administrative Reform Association of 1855". *Victorian Studies* 8.3, pp. 231–242.

ANDERSON, O. (1967). "The Political Uses of History in Mid Nineteenth-Century England". *Past Present* 36, pp. 87–105.

ANEJA, A. and XU, G. (Aug. 2024). "Strengthening State Capacity: Civil Service Reform and Public Sector Performance during the Gilded Age". *American Economic Review* 114.8, pp. 2352–87.

ANZIA, S. F. and TROUNSTINE, J. (2025). "Civil Service Adoption in America: The Political Influence of City Employees". *American Political Science Review* 119.2, pp. 549–565.

ASHRAF, N., BANDIERA, O., DAVENPORT, E., and LEE, S. S. (May 2020). "Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services". en. *American Economic Review* 110.5, pp. 1355–1394.

AYLMER, G. E. (1980). "From Office-Holding to Civil Service: The Genesis of Modern Bureaucracy: The Prothero Lecture". *Transactions of the Royal Historical Society* 30, pp. 91–108.

BESLEY, T., BURGESS, R., KHAN, A., and XU, G. (2022). "Bureaucracy and Development". *Annual Review of Economics* 14.Volume 14, 2022, pp. 397–424.

BESLEY, T. and PERSSON, T. (2011). *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton University Press.

BOSTASHVILI, D. and UJHELYI, G. (2019). "Political budget cycles and the civil service: Evidence from highway spending in US states". *Journal of Public Economics* 175, pp. 17–28.

BOWIE GRANGER, A. (Apr. 1875). "The St. James's Magazine". *St. James's Magazine; London* ser.3. 1, pp. 284–289.

BROWN, G. W. (May 1879). "English Civil Service Reform". *The Atlantic Monthly* 43.258.

CALLAWAY, B. and SANTANNA, P. H. (2021). "Difference-in-Differences with multiple time periods". *Journal of Econometrics* 225.2, pp. 200–230.

CHAISEMARTIN, C. DE and D'HAULTFUILLE, X. (Sept. 2020). "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects". *American Economic Review* 110.9, pp. 2964–96.

CHAPMAN, J. (2025). "Democracy, Redistribution, and Inequality: Evidence from the English Poor Law". Mimeo.

CHAPMAN, R. and GREENAWAY, J. (1980). *The Dynamics of Administrative Reform*. Croom Helm.

CHESTER, D. N. (1981). "The English administrative system 1780-1870".

COLONNELLI, E., PREM, M., and TESO, E. (Oct. 2020). "Patronage and Selection in Public Sector Organizations". en. *American Economic Review* 110.10, pp. 3071–3099.

COMPTON, J. M. (1968). "Open Competition and the Indian Civil Service, 1854-1876". *The English Historical Review* 83.327, pp. 265–284.

CORNELL, A. and SVENSSON, T. (2023). "Colonial origins of modern bureaucracy? India and the professionalization of the British civil service". *Governance* 36.2, pp. 533–553.

DAVIS, J. (2006). "Meritocracy in the Civil Service, 18531970". *The Political Quarterly* 77.s1, pp. 27–35.

DESERRANNO, E. (Jan. 2019). "Financial Incentives as Signals: Experimental Evidence from the Recruitment of Village Promoters in Uganda". *American Economic Journal: Applied Economics* 11.1, pp. 277–317.

DEWEY, C. J. (1973). "The Education of a Ruling Caste: The Indian Civil Service in the Era of Competitive Examination". *The English Historical Review* 88.347, pp. 262–285.
EATON, D. B. (1880). *Civil Service in Great Britain: A History of Abuses and Reforms, and Their Bearing Upon American Politics*. English. New York.
ESTRADA, R. (2019). "Rules versus Discretion in Public Service: Teacher Hiring in Mexico". *Journal of Labor Economics* 37.2, pp. 545–579.
FINAN, F., OLKEN, B., and PANDE, R. (2015). *The Personnel Economics of the State*. en. Tech. rep. w21825. Cambridge, MA: National Bureau of Economic Research, w21825.
FOLKE, O., HIRANO, S., and SNYDER, J. M. (2011). "Patronage and Elections in U.S. States". *American Political Science Review* 105.3, pp. 567–585.
FRY, G. K. (1969). *Statesmen in Disguise: The Changing Role of the Administrative Class of the British Home Civil Service, 1853–1966*. London: Macmillan.
GOODMAN-BACON, A. (2021). "Difference-in-differences with variation in treatment timing". *Journal of Econometrics* 225.2, pp. 254–277.
GRINDLE, M. S. (2012). *Jobs for the Boys*. Harvard University Press.
HANHAM, H. J. (1960). "4. Political Patronage at the Treasury, 18701912". *The Historical Journal* 3.1, pp. 75–84.
HMG, H. M. G. (1870). "Clause V Order in Council 1870".
HUGHES, E. (1942). "HISTORICAL REVISION. No. CI: Civil Service Reform 18535". *History* 27.105, pp. 51–83.
JOHNSON, R. N. and LIBECAP, G. (1994a). *The Federal Civil Service System and the Problem of Bureaucracy*. National Bureau of Economic Research, Inc.
JOHNSON, R. N. and LIBECAP, G. D. (1994b). "Patronage to Merit and Control of the Federal Government Labor Force". *Explorations in Economic History* 31.1, pp. 91–119.
JOHNSTON, M. (1979). "Patrons and Clients, Jobs and Machines: A Case Study of the Uses of Patronage". *American Political Science Review* 73.2, pp. 385–398.
KELLETT, E. E. (1936). *As I Remember*. V. Gollancz, Limited, p. 400.
KETTERING, S. (1988). "The Historical Development of Political Clientelism". *The Journal of Interdisciplinary History* 18.3, pp. 419–447.
KUIPERS, N. and SAHN, A. (2023). "The Representational Consequences of Municipal Civil Service Reform". *American Political Science Review* 117.1, pp. 200–216.
LEUCHT, L. (2023). "Jobs for Votes: Patronage and Performance in Tammany Hall's NYPD". Job Market Paper.
MACDONAGH, O. (1958). "The Nineteenth-Century Revolution in Government: A Reappraisal". *The Historical Journal* 1.1, pp. 52–67.
MOCANU, T. (2023). "Designing Gender Equity: Evidence from Hiring Practices and Committees". Mimeo.
MOORE, R. J. (1964). "The Abolition of Patronage in the Indian Civil Service and the Closure of Haileybury College". *The Historical Journal* 7.2, pp. 246–257.
MOREIRA, D. and PÉREZ, S. (2022). *Who Benefits from Meritocracy?* en. Tech. rep. w30113. Cambridge, MA: National Bureau of Economic Research, w30113.
MOREIRA, D. and PÉREZ, S. (July 2024). "Civil Service Exams and Organizational Performance: Evidence from the Pendleton Act". *American Economic Journal: Applied Economics* 16.3, pp. 250–91.
NORTH, D. C. and WEINGAST, B. R. (1989). "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England". *The Journal of Economic History* 49.4, pp. 803–832.
ORNAGHI, A. (2016). "Civil Service Reforms: Evidence from US Police Departments". Job Market Paper.

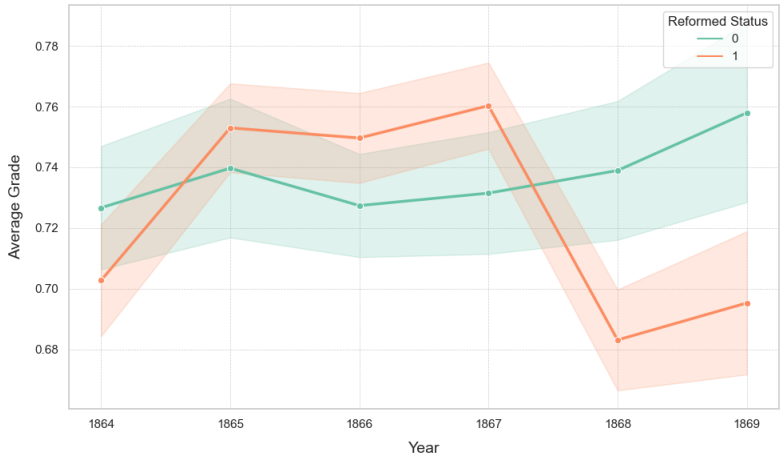
PARLIAMENT OF THE UNITED KINGDOM (Apr. 1869). *Civil And Diplomatic AppointmentsResolution*. House of Commons Debates (Hansard). Commons Chamber, Resolution debated 9 April 1869.
PARLIAMENT OF THE UNITED KINGDOM (June 1870a). *Civil Service Competitive Examinations.Question*. House of Commons Debates, Hansard. Parliamentary Debate. Commons Sitting, 13 June 1870.
PARLIAMENT OF THE UNITED KINGDOM (July 1870b). *Public Service (Competition), *Hansard*, Volume 203: debated on Tuesday 12 July 1870*. Debates of the House of Commons. Parliamentary Debate. UK Parliament, Hansard.
PARLIAMENT OF THE UNITED KINGDOM (Feb. 1870c). *Civil Service Examination Resolution*. House of Commons Debates (Hansard).
PARLIAMENT OF THE UNITED KINGDOM (Mar. 1873a). *Civil Service ExaminationPublic DepartmentsQuestion*. House of Commons Debates (Hansard). Commons Chamber, Volume 215: debated on Monday 31 March 1873.
PARLIAMENT OF THE UNITED KINGDOM (Mar. 1873b). *Questions*. House of Commons Debates (Hansard). Commons Chamber, Volume 215: debated on Tuesday 25 March 1873.
PARRIS, H. (1960). "The Nineteenth-Century Revolution in Government: A Reappraisal Reappraised". *The Historical Journal* 3.1, pp. 17–37.
PHILLIPS, J. A. and WETHERELL, C. (1995). "The Great Reform Act of 1832 and the Political Modernization of England". *The American Historical Review* 100.2, pp. 411–436.
PLUMB, J. and TREVELYAN, G. (1955). *Studies in Social History: A Tribute to G. M. Trevelyan*. Essay index reprint series. Longmans, Green.
RAUCH, J. E. and EVANS, P. B. (2000). "Bureaucratic structure and bureaucratic performance in less developed countries". *Journal of Public Economics* 75.1, pp. 49–71.
RICHARDS, P. G. (1963). *Patronage in British Government*. Toronto: University of Toronto Press.
ROACH, J. (1971). *Public Examinations in England 18501900*. Cambridge Texts and Studies in the History of Education. Cambridge University Press.
ROBINSON, J. and VERDIER, T. (2013). "The Political Economy of Clientelism". *Scandinavian Journal of Economics* 115.2, pp. 260–291.
ROSEVEARE, H. (1973). *The Treasury, 1660-1870: The Foundations of Control*. Vol. 8. Historical Problems. Routledge, p. 216.
SANTANNA, P. H. and ZHAO, J. (2020). "Doubly robust difference-in-differences estimators". *Journal of Econometrics* 219.1, pp. 101–122.
SORAUF, F. J. (1956). "State Patronage in a Rural County". *American Political Science Review* 50.4, pp. 1046–1056.
STANSKY, P. (1973). *The Victorian revolution: government and society in Victoria's Britain. Edited and with an introd. by Peter Stansky*. eng. New York: New Viewpoints.
"The Northcote-Trevelyan Report" (1854).
THERIAULT, S. M. (2003). "Patronage, the Pendleton Act, and the Power of the People". *The Journal of Politics* 65.1, pp. 50–68.
TING, M. M., SNYDERJR, J. M., HIRANO, S., and FOLKE, O. (2013). "Elections and reform: The adoption of civil service systems in the U.S. states". *Journal of Theoretical Politics* 25.3, pp. 363–387.
UJHELYI, G. (2014a). "Civil service reform". *Journal of Public Economics* 118, pp. 15–25.
UJHELYI, G. (May 2014b). "Civil Service Rules and Policy Choices: Evidence from US State Governments". *American Economic Journal: Economic Policy* 6.2, pp. 338–80.
VOTH, H.-J. and XU, G. (2019). "Patronage for Productivity: Selection and Performance in the Age of Sail". *CEPR Discussion Paper* No. 13963.
WEBER, M. (1922). *Economy and Society*. 4th ed. Tübingen.
WHITE, L. D. (1932). "The British Royal Commission on the Civil Service". *The American Political Science Review* 26.2, pp. 315–318.

WILSON, J. Q. (1961). "The Economy of Patronage". *Journal of Political Economy* 69.4, pp. 369–380.
WORLD BANK (2000). *Reforming public institutions and strengthening governance : a World Bank strategy*. English. World Bank e-Library. Washington, DC: World Bank, Public Sector Board, Poverty Reduction and Economic Management.
XU, G. (Nov. 2018). "The Costs of Patronage: Evidence from the British Empire". en. *American Economic Review* 108.11, pp. 3170–3198.
XU, G. (Mar. 2023). "Bureaucratic Representation and State Responsiveness during Times of Crisis: The 1918 Pandemic in India". en. *The Review of Economics and Statistics* 105.2, pp. 482–491.

A Appendix for “Resisting reform in the public sector"

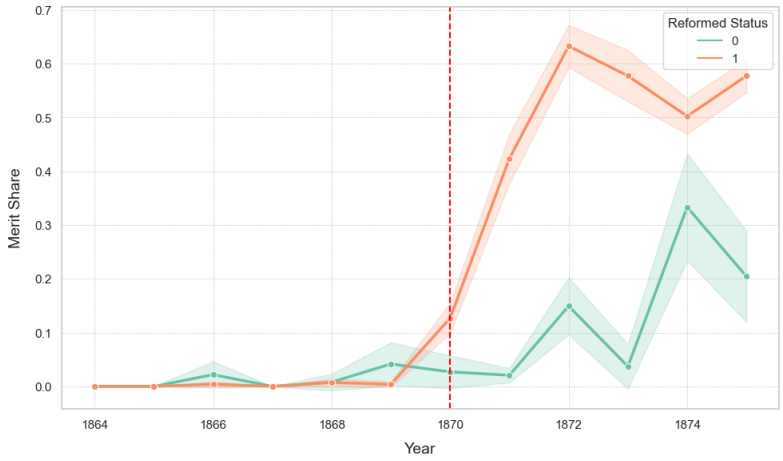
Figures

Figure A2: Average Total Grade pre-Treatment



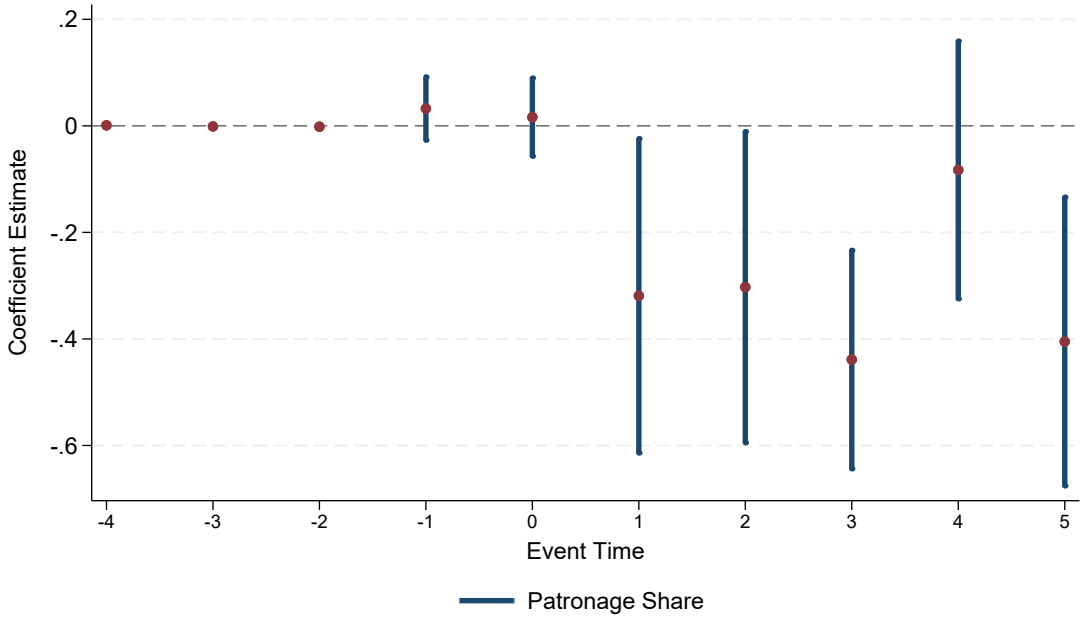
Notes: The figure presents the yearly average of all grades received in reformed and unreformed departments (treated as two groups). This average is calculated from the relative total grade of each applicant, which is the simple ratio of the maximum attainable grade to their achieved grade.

Figure A3: Merit share over period of observation



Notes: The figure presents the annual merit share of entrants in reformed or unreformed departments over the period 1864-1875. This is the inverse of the patronage share, which I will refer to most often in the paper – simply the number of candidates as a percentage of all total that enter through open access competition. The year of reform, 1870, is denoted with a red dashed line in the plot.

Figure A4: Event Study - Conditional DiD Estimates



Notes: The figure above contains an event study made using the CSDiD command in Stata. I condition on the pretreatment hiring share averages (open, limited and direct hiring), allowing the unconditional parallel trend assumption to be relaxed. In the above row I show the ATT in each period using open competition hiring shares as the dependent variables. In the later I show the same using limited competition as the dependent variable. All regressions use 95% confidence intervals.

Tables

Table A5: Reformed Departments

Departments given as reformed in Schedule A
Admiralty; Board of Trade; Charity Commission; Chief Secretary's Office (Ireland); Civil Service Commission; Colonial Office; Constabulary (Ireland); Convict Prisons Department (Ireland); County Courts Judgements Registry; Customs; Dublin Metropolitan Police; Education Office; Emigration Office; Exchequer and Audit Office; India Office; Inland Revenue; Inspector of Reformatories; London Gazette; Lunacy Commission; Metropolitan Police Courts; Mint; National Debt Office; Office of Woods; Office of Works; Petty Sessions Clerks' Registry (Ireland); Poor Law Board; Post Office; Privy Council Office; Queen's and Lord Treasurer's Remembrancer's Office (Scotland); Record Office; Record Office (Ireland); Register Office; Register Office (Ireland); Registry of Deeds (Ireland); Registry of Designs; Stationery Office; Treasury, War Office.

Table A6: Baseline using leads and lags

	Dep. Var: Patronage (1/0)	
	(1)	(2)
Reform × 1870	-0.138	-0.0187
	(0.109)	(0.126)
Reform × 1871	-0.441***	-0.283*
	(0.157)	(0.152)
Reform × 1872	-0.521**	-0.455**
	(0.215)	(0.204)
Reform × 1873	-0.578***	-0.536***
	(0.150)	(0.161)
Reform × 1874	-0.207	-0.289
	(0.192)	(0.191)
Reform × 1875	-0.411*	-0.419*
	(0.241)	(0.229)
Observations	9,309	9,309
Dept FE		✓
p-value: All pre-trends =0	0.411	0.737
Mean Dep. Var.	0.788	0.788

Notes: The regression model is a linear probability model, where I regress a binary variable for reform upon each individual year in the sample, giving the estimated effect for treated departments in each year 1864-1875 (omitting the last pre-treatment year, 1869). I employ fixed effects in column 2 at the department level to minimise individual quirks in the hiring process.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A7: Panel DiD

	Dep. Var: Patronage Share (%)		
	(1)	(2)	(3)
Reform × Post	-0.333***	-0.381***	-0.369***
	(0.0584)	(0.0593)	(0.0590)
Observations	652	605	605
Dept FE		✓	✓
Year FE			✓
Mean Dep. Var.	0.865	0.866	0.866

Notes: The model used is a panel difference in differences, where I regress a treatment indicator (reformed*post) on the share of a departments entrants that enter through patronage. I include a departmental, and annual, fixed effect to deal with heterogeneity at this level.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A8: Conditional DiD Estimates

	Dep. Var: Patronage Share (%)
	(1)
Pre-Treatment Average	0.00773
	(0.00756)
ATT	-0.255***
	(0.0699)
Observations	483
Model	Stable IPW

Notes: The above tables shows the results of the DiD estimator provided by Callaway and SantAnna 2021. This allows one to condition on the pretreatment hiring share averages for each department, improving the comparison between the two and relaxing unconditional parallel trends. Results are reported for the average treatment effect on the treated and the pre-treatment average.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A9: ATT's using matching

Matching Technique:	*Dep. Var:* Patronage Share (%)				
	NN (2)	NN (3)	NN (4)	PS	CEM
	(1)	(2)	(3)	(4)	(5)
Reformed × **Post**	-0.400***	-0.402***	-0.399***	-0.437***	-0.279***
	(0.0452)	(0.0451)	(0.0452)	(0.0396)	(0.0735)
Observations	600	600	600	600	629
Mean Dep. Var.	0.869	0.869	0.869	0.869	0.868

Notes: The above table gives a series of robustness checks done using different matching techniques. The technique used is given by the title at the top of each column. The final column uses CEM matching to weight a simple regression (without fixed effects) of the treatment. All matching is done using pre-treatment hiring behaviour to give comparable treated-control groups for the estimation.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A10: Probit and Logit models

Model:	*Dep. Var:* Patronage (1/0)			
	Probit	Logit	Probit (Jackknife)	Logit (Jackknife)
	(1)	(2)	(3)	(4)
Reformed × **Post**	-1.843***	-3.803***	-1.843***	-3.802***
	(0.515)	(1.199)	(0.190)	(0.532)
Observations	9,309	9,309	9,310	9,310
Mean Dep. Var.	0.788	0.788	0.788	0.788

Notes: The above table gives a series of robustness checks done using probit and logit specifications instead of the baseline LPM model. These techniques use individual data, as in Table 2, but without any fixed effects due to the nature of probit/logit models. The type of model is given in the title at the top of each column – the only change in the latter two columns is the introduction of jackknifed standard errors to further illustrate robustness.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A11: ATT refining treatment/control group

	Dep. Var: Patronage (1/0)		
	(1)	(2)	(3)
Reform $\times$ **Post**	-0.422***	-0.400***	-0.362***
	(0.126)	(0.128)	(0.119)
Observations	9,124	9,087	9,087
Dept FE		✓	✓
Year FE			✓
Mean Dep. Var.	0.786	0.786	0.786

Notes: The regression model is a linear probability model, where I regress the interaction term (reformed * post-treatment) upon a binary variable for patronage use for each individual appointment from 1864-1875 in the British civil service, thus estimating whether reformed departments changed their manner of appointment post reform. I employ fixed effects at the department level to minimise individual quirks in the hiring process. The final step taken in this model is to refine the treated and control groups by dropping three organisations, the Civil Service Commission (which might opt into treatment) and the Foreign and Home Office's, which may have 'opted out' of treatment due their position rather than idiosyncratic bargaining among the executive.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A12: Whitehall and Ancient Departments

Whitehall (**Ancient in bold**)
Admiralty; Board of Trade; Charity Commission; Civil Service Commission; **Colonial Office**; Diplomatic Service; Ecclesiastical Commission; **Education Office**; **Foreign Office**; **Home Office**; House of Commons; India Audit Office; **India Office**; Inspector of Reformatories; Local Government Board; Lunacy Commission; Metropolitan Police; Mint; National Debt Office; Office of Comptroller; Office of Works; Office of Woods; Parliament Office; Parliamentary Counsel; Parliamentary Counsel's Office; Paymaster-General's Office; Privy Council Office; Privy Seal Office; Record Office; Register Office; **Treasury**; **War Office**.

Table A13: ATT's for foreign versus domestic organisations

	Dep. Var: Patronage (1/0)		
	(1)	(2)	(3)
Reformed × **Post** × **Foreign**	-0.181	-0.165	-0.178
	(0.271)	(0.230)	(0.201)
Observations	9,309	9,272	9,272
Dept FE		✓	✓
Year FE			✓
Mean Dep. Var.	0.788	0.787	0.787

Notes: The regression model is a linear probability model, where I regress the interaction term on a binary variable for whether an applicant enters through patronage or not. The coefficient reported is the triple-difference estimate for the ATT difference between foreign and domestic organisations. I employ fixed effects at the department level to minimise department-level quirks in the hiring process. I also use annual fixed effects for a similar reason.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A14: ATT's controlling for positional prestige

Control Var:	*Dep. Var:* Patronage (1/0)			
	1-3 Grouping			All Positions
	(1)	(2)	(3)	(4)
Reformed × **Post**	-0.404***	-0.377***	-0.345***	-0.327**
	(0.126)	(0.130)	(0.121)	(0.141)
Observations	9,309	9,272	9,272	9,267
Dept FE		✓	✓	✓
Year FE			✓	✓
Mean Dep. Var.	0.788	0.787	0.787	0.788

Notes: The regression model is a linear probability model, where I regress the interaction term on a binary variable for whether an applicant enters through patronage or not. Each title gives the control that I use for positional prestige – 1-3 Grouping refers to the groups 1 through 3 that I use for positional prestige (clerk (strict), clerk (weak) and others) while all positions assigns an individual number to each unique title string and then uses this as a control. I employ fixed effects at the department level to minimise department-level quirks in the hiring process. I also use annual fixed effects for a similar reason.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

European
Historical
Economics
Society

EHES Working Paper Series

Recent EHES Working Papers

EHES 287 State Capacity and Subjective Well-Being in Early Modern Spain
Carlos Álvarez-Nogal, Leandro Prados de la Escosura

EHES 286 Build Better Health: Evidence from Ireland on Housing Quality and Mortality
Alan de Bromhead, Ronan C. Lyons, Johann Ohler

EHES 285 Hereditarian Fallacies? Inheritance of Social Status in England, 1600-2022
Gregory Clark

EHES 284 How Long do Wealth Shocks Persist? Less than three generations in England, 1700-2025
Gregory Clark, Neil Cummins

EHES 283 Has oil richness been a force for income equality in Venzuela over the long term?
Pablo Astorga

EHES 282 Female labor force participation in historical census microdata
Jørgen Modalsli

EHES 281 Colonial Governance and Resource Allocation in the British West Indies (1838–1938)
Luisito Bertinelli, Fabio Gatti, Eric Strobl

EHES 280 Why railways fail: Colonial railways and economic development in Habsburg Bosnia-Herzegovina
Magnus Neubert, Stefan Nikolić

EHES 279 Text Analysis Methods for Historical Letters, The case of Michelangelo Buonarrotti
Fabio Gatti, Joel Huesler

All papers may be downloaded free of charge from: https://www.ehes.org/

The European Historical Economics Society is concerned with advancing education in European economic history through study of European economies and economic history. The society is registered with the Charity Commissioners of England and Wales number: 1052680.