Learning Traps

Hassan Sayed*

Princeton University

Current draft: February 4, 2025.
Click here for latest version.

Abstract

Under policy uncertainty, policy choices generate a tradeoff between payoffs and information. Incumbent leaders can implement their preferred policy, but if the outcome is bad, constituents learn this and threaten office removal. I show that, in a repeated setting, the optimal policy of leaders constrained by this threat of information revelation is nonmonotone in beliefs about policy efficacy. At extremal beliefs, leaders experiment with the most extreme policy that keeps them in office. At middle beliefs, they implement spatially intermediate "learning trap" policies that halt learning about the optimal policy when it would be most useful. I apply the model to the post-1861 reforms ending serfdom in Imperial Russia, arguing that by combining elements of liberalism and serfdom to obfuscate inference about policy efficacy, these reforms represent a learning trap, in contrast to the pre-1861 status quo of serfdom and post-1906 experiments with liberal economic reform.

JEL Codes: D72, D73, D78, D83, N33, N40, N43, N53
Word Count: 9,594

*hsayed@princeton.edu . I am deeply indebted to Leah Boustan and Pietro Ortoleva for their diligent work in advising this project, and am especially grateful for Tracy Dennison's and Steven Nafziger's thorough feedback on the Imperial Russian case study. I also thank Roland Bénabou, Dana Foarta, Matias Iaryczower, Gleason Judd, Navin Kartik, Ivan Korolev, Alessandro Lizzeri, Dan McGee, Xiaosheng Mu, Wolfgang Pesendorfer, Kim Sarnoff, Leeat Yariv, Elaine Yao; and participants in the Princeton Microeconomic Theory Student Lunch Seminar, Princeton Political Economy Research Seminar, the 2022 Young Economists Symposium, and the 2023 North American Summer Meeting of the Econometric Society for productive discussions and suggestions. I acknowledge financial support from the William S. Dietrich II Economic Theory Center.

1 Introduction

The policies of incumbent leaders — autocrats, CEOs, or elected officials — are often constrained by threats of being ousted by their constituents — elites, boards of directors, or the electorate. When the optimal policy for constituents is certain, a leader simply implements her most preferred policy that avoids *immediate overthrow*. However, when there is uncertainty about the effects of different policies, policy choices generate both payoffs and information about policy efficacy. If a leader experiments with her preferred policy today and the outcome is good, she can continue to remain in office because constituents will have no incentive to oust her. But if the resulting outcome is bad, constituents may take this as a sign that their preferences are misaligned with their leader's and threaten overthrow, forcing a leader to implement undesirable policies in the future. Uncertainty generates an additional *threat of information revelation* that constrains leaders' willingness to implement certain policies.

This paper develops a repeated model of policy uncertainty where an ideologically-motivated "leader" experiments with different policies, but may be overthrown if she implements policies that harm constituents. It argues that leaders' policy experimentation is nonmonotone in beliefs about whether their preferred policy favors their constituents. At extremal beliefs, when constituents are relatively sure what the optimal policy is, the leader experiments, pursuing extreme policies that prevent overthrow. At intermediate beliefs, when there is greater uncertainty about the optimal policy, the leader halts learning, precisely when it would be most useful. I show that extreme policies — namely, leaders' most and least preferred policies — generate information about policy efficacy, while spatially intermediate "learning trap" policies generate *no information* but modest payoffs for leaders. This creates a tradeoff between information revelation and policy preferences. At high beliefs, information favors leaders, so they pursue their most preferred policy. At intermediate beliefs, information may reveal leaders' preferred policies are harmful, so they implement "learning traps" and forever receive a modest payoff. At low beliefs, constituents threaten overthrow, forcing leaders to pursue their least preferred policies.

I illustrate an application of the model's findings by studying the 1861 economic reforms ending serfdom and codifying the peasant commune in Imperial Russia. Prior to 1861, Russian peasants (serfs) worked noble land as part of communes. Russia's 1856 defeat in the Crimean War represented a shock that generated policy uncertainty, forcing the Tsar to reconsider serfdom. Many in the Tsar's government believed liberal reforms — like private property and freeing labor to move to the cities — could stimulate economic growth and expand its fiscal capacity. However, in the ensuing reforms, the government mixed elements of private and communal property and restricted labor mobility, hampering the economic growth it so desired. I show how this arrangement illustrates the logic of a learning trap: the government feared unrest if liberal policies backfired, but by mixing liberal and communal institutions, distorted labor and production decisions and halted learning about which policy was better. I situate the choice of an intermediate policy amidst policy uncertainty against the two other equilibria of the model: prior to the uncertainty of the 1850s, the government maintained an extreme policy of serfdom, while after shocks favoring liberal reform in 1905, the government experimented with private property.

In my model, a leader prefers a higher policy $x \in [0,1]$. A representative agent for "the people" wants to maximize an outcome $y \in \mathbb{R}$. There are two states of the world: one where the relationship between x and y is positive or single peaked — allowing the leader to implement her preferred policy — and one where it is negative, forcing the leader to implement policies undesirable to her. The leader and people share a common prior. As different policies are implemented, information is revealed. Agents use histories of policy-outcome data (y, x) to infer the state. The key feature is that data on policy efficacy are generated endogenously: by setting x, the leader also controls inference. The people can overthrow the leader at any point, disciplining the policies a leader implements.

Information revolves around a spatially intermediate "learning trap" policy x_{LT} which, when implemented, generates no information about policy efficacy. x_{LT} exists if the people's maximal payoffs in each state of the world are equal. Extreme policies farther from x_{LT} generate more information about the optimal policy, corresponding to policy experimentation.

I show that equilibrium experimentation is nonmonotone in beliefs that the leader's preferred policy benefits the people. At high beliefs, the leader pursues high x, generating favorable flow utility and information. As beliefs decrease, information is less likely to favor her in expectation, forcing her to balance the value of information with policy preferences. Because pursuing x_{LT} obfuscates information, for an intermediate range of beliefs, forever pursuing x_{LT} and achieving a moderate payoff is preferable to pursuing an extreme policy today that risks a lower payoff in the future. This range expands as patience increases, policies become more informative, or the distaste for low x grows. Finally, at low beliefs, the people are certain their preferences are misaligned with the leader's, forcing her to implement low x lest she be overthrown.

In the baseline setting, leaders' policy decisions are the only sources of inference. I extend the model with anticipated and unanticipated *external information revelation*. Unanticipated information — like wars or disasters that "throw back the veil" on policy efficacy — reveals asymmetry in policy variation: leaders switch to experimentation only when large shocks favor their preferred policy. When information is *anticipated*, leaders pursue learning trap policies for a smaller range of beliefs: learning can no longer be completely shut down, so leaders hope experimentation with their preferred policy will favor them. Anticipated shocks can be thought of as political "contagion," where groups of countries with similar historical institutions may learn from each other. I show how these insights can help us understand phenomena like the effect of political fragmentation on differences in experimentation with indirect rule between 19th-century Europe and China.

Reform in Imperial Russia I show how the model can help us understand the tepidness of large-scale economic reform by studying the end of serfdom in Imperial Russia, showing how the Tsar's 1861 emancipation reforms can be seen as a "learning trap" when juxtaposed with pre-1861 serfdom and post-1906 experiments with private property. Prior to 1861, Russian peasants worked state or noble-owned land as members of communes; production and labor choices were managed by seigneurs (Nafziger, 2010). This status quo represents one extreme policy, serfdom. Russia's defeat in the 1853 Crimean War revealed a need to modernize its economy (Starr, 2015). I interpret the war as a shock to beliefs that cast uncertainty about optimal economic policy, marked by heightened discourse about whether serfdom or liberal reform was better. The Tsar's government believed liberal institutions could reduce the power of the nobility (Moon, 2014) jumpstart an industrial revolution (Pereira, 1980), and increase fiscal revenue (Dennison, 2020, 2023). However, it also worried they could backfire, leading to exploitation, abandonment of land, fiscal straits, and overthrow (Dennison, 2014; Polunov et al., 2015).

The model predicts that, in response to uncertainty, the government should shift to an intermediate policy that minimizes learning. Indeed, the Tsar mixed serfdom with Western European institutions like private property. While seigneurs no longer oversaw peasants, communes remained, and many household production, property, and labor allocation decisions required communal consent. Most agricultural land was held communally instead of privately, distorting investment and work incentives (Dennison, 2020). Over many decades, the government used the commune to enforce mobility restrictions, inhibiting urban migration and industrial growth (Nafziger, 2010). I argue that these forces minimized Russians' capacity to discern the benefits of liberal reform. Finally, I interpret political events at the beginning of the 20th century as generating another shock to beliefs in favor of liberal economic institutions. The model predicts that this should cause a radical shift to an extreme policy. Indeed the government began experimenting with private property and other liberal economic institutions as part of Pyotr Stolypin's reforms.

I also address the two traditional explanations of the reform's tepidness: pressures from the nobility and gentry (Khristoforov and Gilley, 2016) and state capacity (Dennison, 2020).

While both forces undoubtedly influenced the reforms, the bargaining power of elites and gentry was relatively weak and slowly waned, and neither force explains why the government did not even gradually reform over four decades. Viewing the reform's intermediacy as a learning trap — in contrast to pre-1861 serfdom and post-1905 reform — provides a fuller picture of its shape in tandem with existing theories.

I introduce the model in section two and solve it in section three. Section four applies the findings of the model to reform in Russia. Section five solves the model with exogenous information revelation with applications to political fragmentation in Europe and China. I conclude in section six. The remainder of the section addresses relevant theoretical literature; literature on Russia is contained in section four and on political fragmentation in section five.

Literature This paper's central insight that leaders pursue obfuscatory policies only at *intermediate* beliefs about policy efficacy contrasts with a political reputation result known as "gambling for resurrection," where leaders pursue their preferred policy at low beliefs and obfuscatory policies at high beliefs (Prendergast and Stole, 1996; Dur, 2001; Majumdar and Mukand, 2004; Fu and Li, 2014; Dewan and Hortala-Vallve, 2019; Tomasi, 2023; Izzo, 2024). This stark difference is driven by the present paper's repeated environment of policy uncertainty. In finite-time settings, leaders pursue their preferred policies in the last period; at high beliefs, obfuscatory policies allow leaders to receive a moderate payoff today and a high payoff tomorrow. However, in repeated settings, leaders must deal with the consequences of their actions for the rest of time. They face a stationary problem where they either choose between a learning trap — perpetually receiving a moderate payoff that never reveals information — or a one-off extreme policy that reveals information. Information revelation will only be preferred to a moderate payoff at high beliefs. Second, leaders in "gambling for resurrection" models can be replaced by an outside option — for example, when there is uncertainty over a politician's ability. [1] At low beliefs, leaders can then only be retained if they experiment and generate a strong signal. By contrast, under policy uncertainty, constituents can threaten overthrow only when they are sure the leader's policy is bad and, even then, the leader can retain office by keeping them indifferent to overthrow. The present paper's insights are conceptually related to partisan traps (de Mesquita and Dziuda, 2024), where voters perpetually elect partisan ideologues who never pursue common value policies; however, "partisan traps" arise due to information asymmetries between politicians and voters that obfuscate the feasibility of non-partisan policies.

[1] The learning trap policy in this paper delivers the same utility in all states of the world. If constituents prefer a learning trap to a (constant) outside option at one belief, they prefer it at all beliefs, meaning the paper's result could not hold with an outside option.

This paper relates to microeconomic studies of reputation in repeated environments. Holmström (1999) shows that managers oversupply effort when firms set future wages based on inferences about their ability, although this simply slows learning.[2] Non-learning often depends on asymmetries in costs (Manso, 2011), information (Ely and Välimäki, 2003), or signal interpretation (Aghion and Jackson, 2016). In the latter, constituents threaten to replace incumbents until a high-ability leader emerges which forces learning in equilibrium; however, this mechanism is driven again by a "gambling for resurrection phenomenon" which cannot be replicated in the present paper. This paper finds more patient leaders prefer to slow learning since the salience of *future* constraints on policies increases, the opposite of Besley and Case (1995) and Banks and Sundaram (1998).

This paper contributes to research on inference using policy-outcome data that argues agents inefficiently learn (Spiegler, 2016; Eliaz and Spiegler, 2020; Levy and Razin, 2021a; Schwartzstein and Sunderam, 2021; Montiel Olea et al., 2022). These data may hamper learning about policies or politician ability. Voters learning from spatial platforms may converge on inefficient policies (Callander, 2011). Frequentist inference or bounded memory can cause polarization (Levy and Razin, 2021b; Izzo et al., 2021). Less information about policies can be better for voters (Prat, 2005) or politicians (Kartik et al., 2015). Substitutability between politician effort and ability (Ashworth et al., 2017) can hamper learning about ability. Empirical work also documents inferences about policy efficacy, including trust in government (Chen and Yang, 2019), external validity of policy experiments (Wang and Yang, 2021), and policy contagion (Buera et al., 2011; Mukand and Rodrik, 2005).

A common theme throughout models of autocratic power-balancing is that rulers self-regulate their exploitation in order to avoid overthrow; they implement the most extreme policies keeping constituents or elites indifferent to overthrowing them (Acemoglu and Robinson, 2000, 2001; De Mesquita and Smith, 2010; Boix and Svolik, 2013; Li et al., 2014; Dower et al., 2018).[3] This paper argues that leaders *moderate* their actions when policies generate information. It also relates to gradualism in "divide-the-dollar" settings like land reform[4], as discussed in Roland (2002) and Acemoglu and Robinson (2008), showing that under uncertainty about the optimal form of land ownership, "partial reforms" mixing different property regimes can obfuscate which reform is better.

[2]Non-learning in their paper is driven by an evolving state.

[3]The last paper looks at authority given to peasant communes given the frequency of disruptions in 1860s Russia, the period of this paper's case study.

[4]These considerations are important in the context of the Russian case study application, and a microfoundation addressing their specific role is constructed in the online appendix.

2 Model

Policies Consider an infinite-horizon model beginning at $t = 1$. Each period, a policy $x_t \in X = [0,1]$ is implemented which determines an outcome $y_t \in \mathbb{R}$ realized later that period. For example, y can represent profits, income, or the value of a public good.

Agents The first agent in the model is the "leader" — an entrenched incumbent, autocrat, or CEO. The second agent — "the people" — constrains the leader's power and can be interpreted as a representative agent for an electorate, board of directors, or elite group. Both agents are infinitely lived.

The leader discounts the future at rate δ. Her flow utility is $(1-\delta)u_\ell(x_t)$ for $u_\ell(\cdot)$ strictly increasing, differentiable, and weakly concave.[5] Preference for higher x_t could reflect an ideological bent; reputational concerns; or the size of rents the government is able to extract in a setting like land reform.[6]

The people's flow utility is y_t. They discount the future at rate 0. This assumption simplifies the model to illuminate its key tradeoffs, but the insights are identical with nonmyopia or other frictions, studied in section 3.2.

Actions At $t = 1$, the leader holds power. At the beginning of each period the leader is in power, the people decide whether to incur a one-time cost $c > 0$ to overthrow the leader. If they overthrow, the people set x_t from t onwards. Overthrow succeeds with probability 1; the leader receives a loss $-L < u_\ell(0)$ and no further flow utility. If the leader is not overthrown, she sets x_t for that period and t moves to $t+1$.

States Agents possess uncertainty over two states of the world governing the relationship between x and y. In the "pro-leader" state, increases in x cause increases in y up to a single-peak $1 \geq \tilde{x} > 0$. In the "anti-leader" state, increases in x cause decreases in y:

$$y_t = -(x_t - \tilde{x})^2 + \epsilon_t \equiv f_g(x_t) + \epsilon_t \qquad \text{Pro-Leader}$$
$$y_t = -x_t^2 + \epsilon_t \equiv f_b(x_t) + \epsilon_t \qquad \text{Anti-Leader}$$

[5] We assume concavity for tractability; all results go through with $u_\ell(\cdot)$ strictly increasing using a piecewise concavification of $u(\cdot)$.

[6] The "Additional Results" section of the online appendix explores a fiscal microfoundation in the setting of the paper's Russian case study. Specifically, x represents both the degree of involvement of the gentry in intermediating taxation and access to private property. Increasing x reduces the degree of gentry intermediation in revenue collection, improving the share the government can keep for itself.

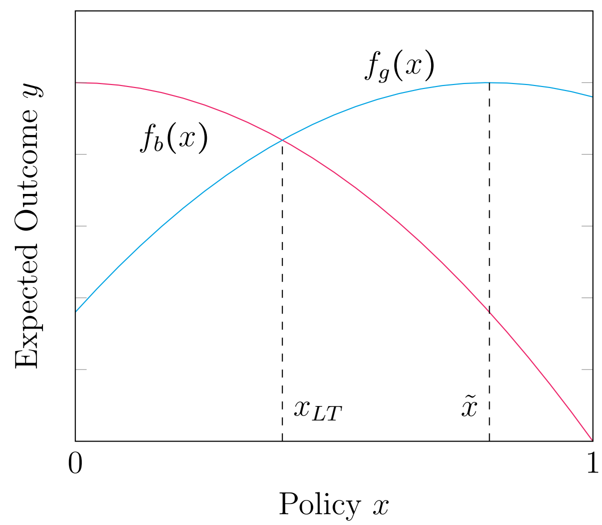

Figure 1: $f_g(x)$ and $f_b(x)$ graphed

$\epsilon_t \sim \mathcal{U}[-\sigma, \sigma]$ is i.i.d. every period, making inference potentially imperfect; Proposition 7 of the "Additional Results" appendix shows a version of the paper's main result for general distributions of ϵ_t. f_g and f_b are graphed in Figure 1. [7]

The expected value of y conditional on $x = \tilde{x}$ is higher in the pro-leader than the anti-leader state, and $x = 0$ is better in the anti-leader than the pro-leader state. In particular this means that $f_g(x) - f_b(x)$ is single crossing at some $x_{LT} \in (0,1)$ with $y_{LT} \equiv f_g(x_{LT})$.[8]

Inference Both agents follow Bayes' Rule and use histories $\{y_\tau, x_\tau\}_{\tau \leq t}$ to update a common belief over the two states. Denote q_{t-1} the belief in the pro-leader state at the beginning of time t. During period t, agents observe a policy-outcome pair (y_t, x_t) and use it to update q_{t-1} to q_t, to be carried into $t+1$. Denote $\Delta(x) = \min\{|f_g(x) - f_b(x)|, 2\sigma\}$ the *effective difference* between the expectation of y conditional on x in each state. Bayes' Rule implies that the distribution of posteriors conditional on x_t and q_{t-1} is:

$$q_t | q_{t-1}, x_t = \begin{cases} 1 & \text{with prob. } \frac{\Delta(x_t)}{2\sigma} q_{t-1} \\ q_{t-1} & \text{with prob. } 1 - \frac{\Delta(x_t)}{2\sigma} \\ 0 & \text{with prob. } \frac{\Delta(x_t)}{2\sigma}(1 - q_{t-1}) \end{cases}$$

When $x_t = x_{LT}$, $f_g = f_b$, meaning $\Delta(x_{LT}) = 0$. x_{LT} shuts down information revelation

[7]All results hold if f_g and f_b are linear with positive and negative slope, respectively; as well as under moderate asymmetry, which we analyze later.

[8]We can also allow u_ℓ to depend on the state of the world; for the main results to go through, we simply need that the utility in the anti-leader state of the world is worse than x_{LT}, which is worse than in the pro-leader state of the world.

($q_{t-1} = q_t$ with probability 1) and we refer to it as a **learning trap policy**. [9] While extreme policies reveal more information and are preferable when uncertainty is resolved, x_{LT} is a "muddied policy" that mixes elements of these opposing extremes to hamper inference. Finally, we assume $c < x_{LT}^2$ so that the overthrow threat binds in equilibrium.

Sequential Game We describe the game recursively.

1. At time t, the people decide whether to overthrow the leader.
   (a) If the people decide to overthrow the leader:
      i. The people incur the cost of overthrow c. The leader loses power, receives $-L$, and receives no further flow utility.
      ii. The people set x_t, y_t realizes, and they receive utility y_t.
      iii. The people update q_{t-1} to q_t. Return to step 1(a)ii at $t + 1$.
   (b) If the people decide not to overthrow the leader:
      i. The leader sets x_t, y_t realizes, and the people receive utility y_t. The leader in power receives $(1 - \delta)u_\ell(x_t)$.
      ii. Agents update their belief q_{t-1} to q_t. Returns to step 1 at $t + 1$.

We focus on pure-strategy Markov Perfect Equilibria with respect to the belief q_{t-1}. $x_\ell(q) \in [0, 1]$ indicates the leader's policy plan. The people's strategy is given by $\{p(q), x_p(q)\}$. $p(q) \in \{\text{overthrow}, \text{nooverthrow}\}$ denotes the overthrow decision. $x_p(q) \in [0, 1]$ indicates their policy plan once in power. $x_\ell(q)$ and $\{p(q), x_p(q)\}$ must best respond to each other in equilibrium.

3 Analysis

We first analyze the baseline model, highlighting the forces predicting nonmonotonic experimentation and policy moderation at intermediate beliefs. We then relax the people's myopia and introduce costly experimentation to illustrate how the model's insights generalize.

[9] The existence of x_{LT} only depends on the single-crossing property, and generalizes beyond uniform noise. If $\epsilon_t \sim p(\epsilon_t)$ for some distribution p :

$$q_t = \frac{p(y_t - f_g(x_{LT}))q_{t-1}}{p(y_t - f_g(x_{LT}))q_{t-1} + p(y_t - f_b(x_{LT}))(1 - q_{t-1})} = \frac{p(y_t - f_g(x_{LT}))q_{t-1}}{p(y_t - f_g(x_{LT}))q_{t-1} + p(y_t - f_g(x_{LT}))(1 - q_{t-1})}$$
$$= \frac{p(y_t - f_g(x_{LT}))q_{t-1}}{p(y_t - f_g(x_{LT}))} = q_{t-1}$$

3.1 Analysis of Baseline Model

People The people's optimal policy, conditional on being in power, is $x_p(q) = \arg\max_x \mathbb{E}[y|x]$; expectations are taken with respect to q. The people have a strict incentive to overthrow if and only if

$$\max_x \mathbb{E}[y|x] - c > \mathbb{E}[y|x_\ell(q)],$$

i.e. if the best policy they can achieve is better than the leader's, less the cost of overthrow; in equilibrium, they overthrow if and only if the inequality above is satisfied strictly. For each q, we define $\text{NR}(q)$ as the set of policies that weakly prevents overthrow: $\text{NR}(q) = \{x_\ell(q) \in [0,1] : \max_x \mathbb{E}[y|x] - c \leq \mathbb{E}[y|x_\ell(q)]\}$. $\text{NR}(q)$ is monotone in q, in the sense that its minimal and maximal elements are increasing in q. Because x_{LT} delivers the same utility in both states of the world, there exist a range of interior beliefs where x_{LT} is contained in the interior of $\text{NR}(q)$. Its graph as a function of q is shaded below.

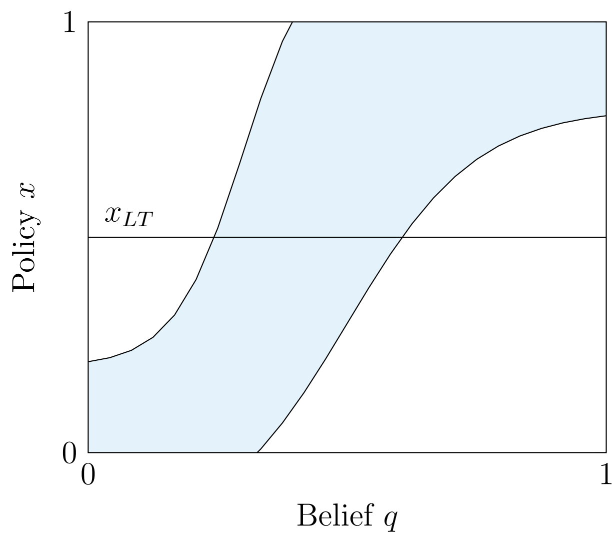

Figure 2: No-Overthrow Constraint $\text{NR}(q)$

Leader's Solution Denote by $\overline{x}(q) \equiv \max\{\text{NR}(q)\}$ the most extreme policy avoiding overthrow. At $q = 0$ or 1, policies have no information externalities, so the leader simply plays this extreme policy for all time. Denote $u_\ell(\overline{x}(0)) = \underline{u}$ and $u_\ell(\overline{x}(1)) = \overline{u}$, which hence represent the value of the leader's problem at these two points. Denote $u_\ell(x_{LT}) = u_{LT}$ and note that, by the assumptions on c, $\underline{u} < u_{LT} < \overline{u}$. This means that the leader's utility at $q = 1$ is preferred to implementing x_{LT} for all time, but that implementing x_{LT} for all time is *better* than being stuck at $q = 0$.

Since L is large, the leader will always want to avoid overthrow on the equilibrium path. This means that, for interior q, the Bellman Equation describing $x_\ell(q)$ is:

$$V(q) = \max_{x_\ell(q)\in \mathrm{NR}(q)} \underbrace{(1-\delta)u_\ell(x)}_{\text{Flow utility}} + \delta\left(\underbrace{\frac{\Delta(x)}{2\sigma}\Psi(q)}_{\text{Truth revealed}} + \underbrace{(1 - \frac{\delta\Delta(x)}{2\sigma})V(q)}_{\text{Truth not revealed}} \right),$$

where $\Psi(q, x') = q\overline{u} + (1-q)\underline{u}$ is the *expected value from revealing the truth.* When the leader implements a policy x, this policy potentially generates information, in addition to flow utility. With probability $q\Delta(x)/2\sigma$, $q \to 1$. With probability $(1-q)\Delta(x)/2\sigma$, $q \to 0$. In the former case, the leader's value function is $\overline{u}$ and in the latter it is $\underline{u}$. Otherwise, q is invariant.

The following proposition characterizes $x_\ell(q)$ for patient leaders.

Proposition 1. *Suppose $\sigma(1-\delta)$ is small or $\underline{u}$ is sufficiently low. There exists a threshold $\underline{q} \in (0,1)$ such that:*

1. *If $q \leq \underline{q}$, $x_\ell(q) = \min\{x_{LT}, \overline{x}(q)\}$. The leader plays the learning trap or the policy closest to it.*

2. *If $q > \underline{q}$, $x_\ell(q) = \overline{x}(q)$. The leader plays the maximal policy preventing overthrow.*

Proof. For $q \in (0,1)$, consider a problem where the leader is constrained at $q = 0$ or 1 but can otherwise implement any policy in $[\underline{x}(0), \overline{x}(1)]$, which can be rewritten as:

$$\max_{x\in[\underline{x}(0),\overline{x}(1)]} \frac{2\sigma(1-\delta)}{2\sigma(1-\delta)+\delta\Delta(x)}u_\ell(x) + \frac{\delta\Delta(x)}{2\sigma(1-\delta)+\delta\Delta(x)}\Psi(q)$$

This expression is a convex combination of flow utility $u_\ell(x)$ and the value $\Psi(q)$ of the truth. If the leader pursues a strategy where she implements x_{LT} for all time, it delivers a value $u_\ell(x_{LT}) = u_{LT}$. As x moves away from x_{LT}, weight shifts away from flow utility $u_\ell(x)$ and onto the expected value of revealing information $\Psi(q)$.

Suppose q is low, so that $\Psi(q) < u_{LT}$. If the leader implements $x > x_{LT}$, this results in higher flow utility relative to u_{LT}. But it also reveals information, whose expected value $\Psi(q)$ is worse than u_{LT}. If $\sigma(1-\delta)$ is small, the effect of flow utility is small, meaning that pursuing the learning trap for all time and receiving a modest payoff is better than gaining a bit of flow utility that reveals unfavorable information, up to a point $\underline{q}$.

If $\Psi(q) \geq u_{LT}$, information is better in expectation than shutting down learning and receiving a moderate payoff. Pursuing $x > x_{LT}$ both increases flow utility and the weight on the revelation of favorable information. For high q, because information is favorable, $x < x_{LT}$

may counterintuitively be preferred to x_{LT} after a point $\overline{q}$; while this edge case does not bind in the baseline, it may in a more general setting.

The solution to this *unconstrained* problem is shown in the left panel of Figure 3 in red. In particular, we normalize $\overline{u} = 1$, $u_{LT} = x_{LT}$, and let $\underline{u} < 0$ so that $\Psi(q)$ is below 0 for low q. This panel also illustrates the thresholds $\underline{q}$ and $\overline{q}$. Below $\underline{q}$, the leader prefers the learning trap for both flow utility and informational purposes; between $\underline{q}$ and $\overline{q}$, prefers a high policy for flow utility purposes; and above $\overline{q}$, prefers any extreme policy to x_{LT} for informational purposes.

The right panel transposes the constraint set $\mathrm{NR}(q)$ (shaded) onto the utility space of the leader. By overlaying the forces from the unconstrained problem onto this constraint set, we are able to trace out the leader's solution $x_\ell(q)$ in the constrained problem (bold), which forces the leader to pursue $x < x_{LT}$ for low q.

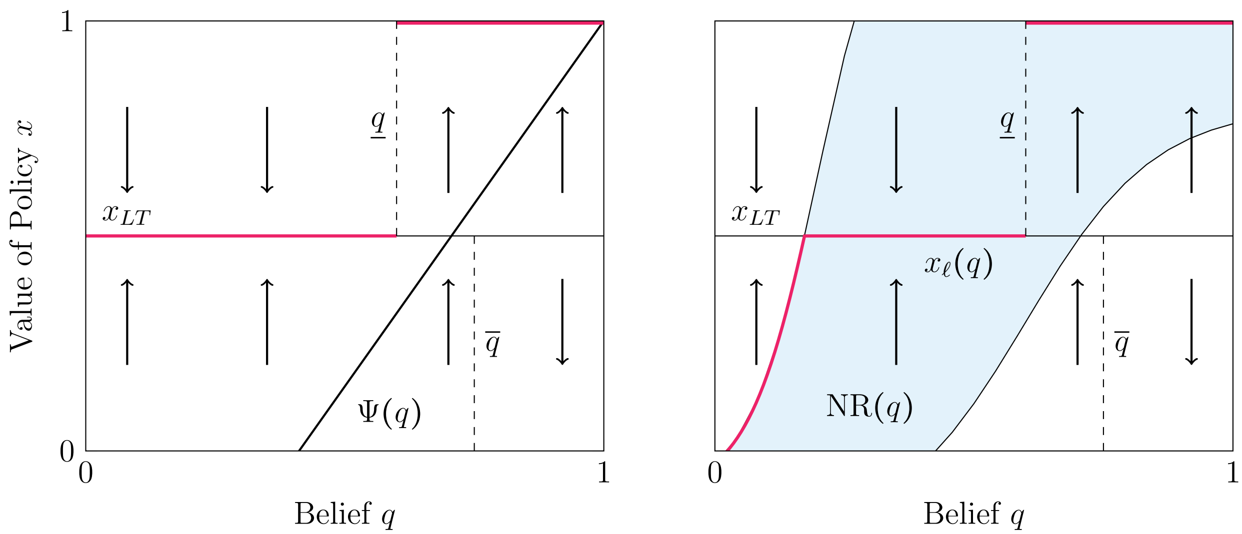

Figure 3: Forces Driving Optimal Policy in Unconstrained Leader's Problem Problem (Left) vs. Constrained Leader's Problem (Right)

□

The result also yields the following comparative statics.

Proposition 2. *Suppose σ decreases, δ increases for the leader, or $\underline{u}$ decreases. Then $\underline{q}$ increases.*

As σ decreases, for each x, the probability of information revelation increases. As δ increases, the leader deemphasizes flow utility.[10] In both cases, the leader is worried more

[10] $(1 - \delta)$ could represent the length of the leader's office tenure and the frequency of evaluations.

about how her actions may reveal the truth, increasing attraction to learning traps. Finally, as $\underline{u}$ decreases, for all q, the value of revealing information decreases.

Comments Proposition 1 highlights a nonmonotonicity in policy experimentation. When the optimal policy is relatively certain, leaders pursue extreme policies. For high q, they implement their most preferred policies, and for low q, the highest policy keeping the people indifferent to overthrow. If the optimal policy is uncertain, leaders moderate policies to tide a threat of information leading to a *future* threat of overthrow — precisely when learning about the optimal policy would be most valuable. Notably, the people have a strict incentive to *retain* the leader at intermediate q, a contrast to dictator games where leaders always keep the people indifferent to overthrow.

A repeated setting is crucial for the result. In a two-period model, a leader would always implement her preferred policy in the second period, meaning she would be overthrown only at low q. This would generate the opposite effect: leaders would implement x_{LT} at high q so as not to "rock the boat," and $x = 1$ at low q to "gamble for resurrection." The result would also not hold if the people received a constant outside option upon overthrow, such as in a setting with uncertainty over a leader's skill. Suppose y_{LT} were preferred to the outside option at $q = 1/2$. At $q = 0$, y_{LT} would still be preferred to the outside option. The leader could then achieve utility at least u_{LT} at $q = 0$ by repeatedly implementing x_{LT}, eroding the need to implement x_{LT} at higher beliefs.

If $\text{NR}(q)$ is increasing in q and its graph convex, a version of this proposition can be shown for arbitrary distributions of ϵ; this is addressed in Proposition 7 of the "Additional Results" online appendix. In practice, and particularly without myopia, this set may not be convex or even connected, meaning standard solution techniques from dynamic programming cannot be used to analyze the problem.

3.2 Model with Nonmyopia and Adjustment Costs

We now allow the people in our model to be infinitely lived and discount the future at rate δ. The leader discounts the future at rate δ^ℓ. The people bear adjustment costs $\kappa(x_t, x_{t-1}) = \kappa|x_t - x_{t-1}|$, capturing frictions from institutional reform or costly experimentation.[11] Their utility is $u_p(y_t, x_t, x_{t-1}) = (1-\delta)y_t - (1-\delta)\kappa|x_t - x_{t-1}|$. Let $k \equiv (1-\delta)\kappa$.

[11] A model where adjustment costs do not depend on the size of an adjustment delivers identical results. Myopia, risk aversion, or replacing the leader with one who wants to minimize x_t could be used as alternative frictions. Forcing the leader to bear costs only strengthens the result.

We allow for asymmetry in the value of y_t between states:

$$y_t = \beta_g - (x_t - \tilde{x})^2 + \epsilon_t \equiv f_g(x_t) + \epsilon_t \qquad \text{Pro-Leader}$$
$$y_t = \beta_b - x_t^2 + \epsilon_t \equiv f_b(x_t) + \epsilon_t \qquad \text{Anti-Leader}$$

We assume:

1. $2k > |f_g'(x_{LT})|, |f_b'(x_{LT})| > k$;

2. $|(\beta_g - \beta_b)/2\tilde{x}| < k$;

3. $\tilde{x}^2/4 > 3|\beta_b - \beta_g| + 3(\beta_b - \beta_g)^2/\tilde{x}^2$;

4. $k \in \left(\frac{2\tilde{x} - \sqrt{\tilde{x}^2/4 - 3|\beta_b - \beta_g| - 3(\beta_b - \beta_g)^2/\tilde{x}^2}}{3}, \frac{2\tilde{x} + \sqrt{\tilde{x}^2/4 - 3|\beta_b - \beta_g| - 3(\beta_b - \beta_g)^2/\tilde{x}^2}}{3}\right)$.

These assumptions bound payoff asymmetry between the two states, and suggest policy adjustment should not be *so* costly that policy never moves, but not so cheap that the people excessively prefer experimentation. When $\beta_b = \beta_g$, this assumption reduces to $2k > \tilde{x} > 5\tilde{x}/6 > k$.

Markov Perfect Equilibrium is calculated with respect to the prior q and the previous period's status quo x'. We write $x_p(q, x')$ as the people's optimal policy conditional on overthrow. Full proofs for this section are provided in the "Proofs of Main Results" part of the online appendix.

Solution to People's Problem We first solve for $x_p(q, x')$. We then use this to characterize the overthrow decision and $x_\ell(q, x')$. $x_p(q, x')$ is the policy function of the following Bellman equation:

$$W(q, x') = \max_{x \in [0,1]} \underbrace{(1-\delta)\mathbb{E}[y|x] - k|x - x'|}_{\text{Flow utility}} + \delta\Bigg(\underbrace{\frac{\Delta(x)}{2\sigma}\big(qW(1,x) + (1-q)W(0,x)\big)}_{\text{Truth revealed}} + \underbrace{\Big(1 - \frac{\Delta(x)}{2\sigma}\Big)W(q,x)}_{\text{Truth not revealed}}\Bigg)$$

The first term represents expected flow utility $qf_g(x) + (1-q)f_b(x)$ less costs $k|x - x'|$. The next terms represent continuation utility. With probability $\frac{\Delta(x)}{2\sigma}$, the truth is revealed. Conditional on this, the belief moves to 1 with probability q and 0 with probability $1-q$. With probability $1 - \frac{\Delta(x)}{2\sigma}$, beliefs remains at q; Today's x becomes tomorrow's status quo policy.

The following proposition describes the value function associated with $x_p(q, x')$, which captures a tradeoff between information and policy flexibility.

Proposition 3. *$x_p(q, x')$ is the solution to the following Bellman equation:*

$$W(q, x') = \max_{x \in [0,1]} \frac{2\sigma(1-\delta)\mathbb{E}[y|x] + \delta\Delta(x)\Phi(q, x)}{2\sigma(1-\delta) + \delta\Delta(x)} - k|x - x'|$$

where $\Phi(q, x) = qW(1, x) + (1-q)W(0, x)$. For intermediate values of c, there exists a nonempty open set $LT \subseteq [0,1] \times X$ with $LT \ni (1/2, x_{LT})$ such that for all $(q, x') \in LT$, $W(q, x') \leq y_{LT} + c \leq W(0, x'), W(1, x')$.

As in the leader's problem in our benchmark, the people's utility is a convex combination of flow utility $\mathbb{E}[y|x]$ and revealing the truth $\Phi(q, x)$, less adjustment costs. Implementing the learning trap x_{LT} generates a constant value y_{LT}.

For intermediate q, the people face a tradeoff: extreme x generates information, but because there is uncertainty about whether high or low x is optimal, they may occur excessive adjustment costs by experimenting in one direction, learning that the optimal policy is in another direction, and then reversing course. This tradeoff means that the difference between $W(q, x')$ and the value of shutting down information y_{LT} is smaller at intermediate beliefs. At extremal q, there is no tradeoff: there is relative certainty about the effects of policies, meaning backtracking costs are not likely to be incurred in the future.

The last part of the proposition establishes bounds on c such that, at $q = 1/2$, if the leader pursues x_{LT}, she strictly prevents overthrow, while at $q = 0$ or 1, she cannot implement x_{LT} without being overthrown. This result suggests intuitive comparative statics: increasing policy adjustment or overthrow costs expands the set of beliefs and status quo policies at which it is not credible to overthrow a leader continually implementing x_{LT}.

Proposition 4. *Let $LT(c, k)$ parametrize the set of values which the learning trap is permissible. Fixing all other parameters, suppose either c or k increases to c' or k'. Then, $LT(c, k) \subsetneq LT(c', k)$ and $\subsetneq LT(c, k')$.*

Overthrow Decision Define $\tilde{W}_{x_\ell(q,x')}$ recursively as follows:

$$\begin{aligned}\tilde{W}_{x_\ell(q,x')}(q, x') = (1-\delta)\mathbb{E}[y|x_\ell(q, x')] - k|x_\ell(q, x') - x'| + \frac{\delta\Delta(x)}{2\sigma}\tilde{\Phi}_{x_\ell(q,x')}(q, x_\ell(q, x')) \\ + \delta(1 - \frac{\Delta(x)}{2\sigma})\tilde{W}_{x_\ell(q,x')}(q, x_\ell(q, x'))\\ \tilde{\Phi}_{x_\ell(q,x')}(q, x) = q\tilde{W}(1, x_\ell(1, x)) + (1-q)\tilde{W}(1, x_\ell(0, x))\end{aligned}$$

This represents the *people*'s value from accepting a leader's policy plan $x_\ell(q,x')$. Lemmata 7 and 8 in the online appendix provide a closed form for this expression and show that, in equilibrium, this strategy has the property that $x_\ell(q,x) = x_\ell(q,x_\ell(q,x))$. If the leader's strategy is only to implement x_{LT} for all time, the value of this objective is $y_{LT} - k|x_{LT} - x'|$.

Let $\text{NR}(q,x')$ be the set of policies that weakly prevents credible overthrow:

$$\text{NR}(q,x') = \left\{x \in [0,1] : \tilde{W}_x(q,x') \geq W(q,x') - c\right\}.$$

The people overthrow the leader if and only if $x_\ell(q,x') \notin \text{NR}(q,x')$.

Leader's Problem Define $\underline{x}_{LT}(q,x')$, $\overline{x}_{LT}(q,x')$, $\underline{x}(q,x')$, and $\overline{x}(q,x')$ as follows:

$$\underline{x}_{LT}(q,x') = \max\{x \in \text{NR}(q,x') : x \leq x_{LT}\} \qquad \overline{x}_{LT}(q,x') = \min\{x \in \text{NR}(q,x') : x \geq x_{LT}\}$$
$$\underline{x}(q,x') = \min\{x \in \text{NR}(q,x')\} \qquad \overline{x}(q,x') = \max\{x \in \text{NR}(q,x')\}$$

$\underline{x}_{LT}$ and $\overline{x}_{LT}$ correspond to the *closest* policies to the learning trap preventing overthrow. $\overline{x}$ and $\underline{x}$ are the lowest and highest policies preventing overthrow altogether.

Precisely as in the benchmark, at $q = 0$ or 1, the leader will pursue the most extreme policy preventing overthrow, i.e. $\overline{x}(q,x')$. Using the assumption on c, we again have:

$$\underline{u}(x') \equiv u_\ell(x_\ell(0,x')) < u_\ell(y_{LT}) < u_\ell(x_\ell(1,x')) \equiv \overline{u}(x')$$

The Bellman describing $x_\ell(q,x')$ is then:

$$V(q,x') = \max_{x_\ell(q,x') \in \text{NR}(q,x')} (1-\delta^\ell)u_\ell(x) + \frac{\delta^\ell \Delta(x)}{2\sigma}\Psi(q,x) + \delta\left(1 - \frac{\delta^\ell \Delta(x)}{2\sigma}\right)V(q,x)$$

where $\Psi(q,x') = q\overline{u}(x') + (1-q)\underline{u}(x')$. The following theorem characterizes $x_\ell(q,x)$ for patient leaders: the leader is either attracted to policies as close to x_{LT} as possible — minimizing information revelation — or an experimental policy that reveals information.

Theorem 1. *Suppose $\sigma(1-\delta^\ell)$ is small or $\underline{u}(x')$ is sufficiently low. Then, there exist thresholds $\underline{q} < \overline{q} \in (0,1)$ such that the following hold:*

1. *If $q \leq \underline{q}$, then $x_\ell(q,x') = \underline{x}_{LT}(q,x')$ or $\overline{x}_{LT}(q,x')$. The leader plays the* closest policy *to the learning trap.*

2. *If $q \in [\underline{q},\overline{q}]$, then $x_\ell(q,x') = \underline{x}_{LT}(q,x')$ or $\overline{x}(q,x')$. The leader plays either the lowest policy below the learning trap or her highest feasible policy.*

3. *If* $q > \overline{q}$*, = then* $x_\ell(q, x') = \overline{x}(q, x')$ *or* $\underline{x}(q, x')$ *(and it is necessary that* $\Delta(\underline{x}(q, x')) > \Delta(\overline{x}(q, x'))$*). The leader either plays her highest feasible policy or policies maximizing information revelation.*

This general formulation shows that the central forces of the model are identical to the baseline: leaders prefer learning traps at intermediate and low beliefs and experiment at high beliefs. At some beliefs, an inability to pursue learning traps without overthrow may generate additional nonmonotonicities. Comparative statics with respect to σ, δ^ℓ, and $\underline{u}(x')$ follow as in the baseline.

3.3 Unanticipated Information Shocks

Comparative statics with respect to q illustrate the effect of *unanticipated* shocks to beliefs in the model — which may represent sudden events like earthquakes, pandemics, or conflicts that test institutions and force leaders to experiment, thereby revealing information. Figure 4 shows how shocks to q at the end of every period affect policy when the status quo is $x' = x_{LT}$, shown below in Figure 4 as q_0 moves either to q_1, q_2, or q_3. With a slight abuse of notation, we write $x(q) = x_\ell(q, x_{LT})$.

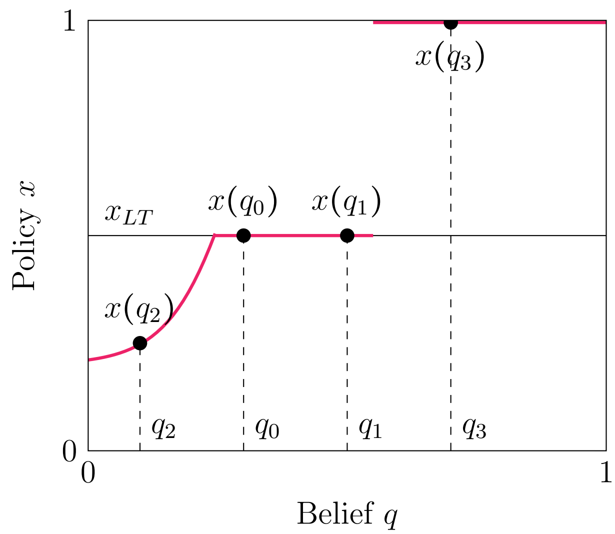

Figure 4: Effect of Unanticipated Shock to q on Policy

If q_0 moves slightly to the right to q_1, there is no policy change even though this information favors the leader. If q_0 drops to q_2, the leader is still attracted to x_{LT}, but must play a lower policy to prevent overthrow. If q_0 jumps to q_3, policy experiences a discontinuous jump from x_{LT} to 1, since information revelation now swings in the leader's favor. Hence, beginning at intermediate beliefs, policy variation is minimal under smaller or negative shocks to low or intermediate beliefs, but discontinuously shifts to experimentation when large shocks favor the leader.

4 Application: The 1861 Reforms as a Learning Trap in Imperial Russia

This section illustrates how the model's insights can be used to understand large-scale economic reform by applying its insights to the institutionalization of the peasant commune in Late Imperial Russia. Prior to 1861, the Russian economy operated under "serfdom": an agricultural land tenure system where peasant farmers worked nobles' land as part of communes overseen by a seigneur. However, faced with a need to modernize its economy in the 1860s, the Russian government considered liberal, Western European style economic reform. By instituting private property rights and allowing peasants freedom to move from their lands, a successful liberal reform would increase the government's fiscal revenue, accelerate Russia's industrial revolution, and encourage economic growth. However, it worried liberal reform was inappropriate to Russia's historical institutions and could backfire, fiscally damaging the Empire and leading to its overthrow.

I argue that, in response to this policy uncertainty, the Russian government pursued a middle ground that mixed elements of both serfdom and liberal institutions. I argue that this intermediate policy can be seen as a "learning trap" that, by mixing these two extremes, obfuscated information about the effects of liberal institutions to tide threats of overthrow. I contrast this learning trap with two other equilibria of the model corresponding to policy experimentation under extremal beliefs — serfdom pre-1861 and experiments with liberal reform after 1905.

Model Mapping To map the model to Imperial Russia, we identify policies and players. One extreme policy is serfdom: The other is a liberal, private property system where peasants own their own land and control individual labor and production decisions. The "people" are a combination of the Russian peasantry and elites. Outcomes y represent aggregate economic wellbeing. The "leader" is the Russian Tsar's government and, for *fiscal* reasons detailed in section 4.1, prefers liberal policies.

In one state of the world, liberal reform is better for Russia's economy than communal serfdom; endowing peasants with land and allowing freedom of movement and production decisions is better than serfdom. The leader's utility is higher: liberal reform increases fiscal revenue. In the other state, the government would be better off under serfdom, concentrating administrative power with elites and limiting freedom, since liberal reform would spread the government thin between compensating the gentry for land expropriation and funding state capacity. The government's utility — measured as fiscal revenue — would be worse under serfdom than under successful liberal reform.

The first subsection describes how a pre-1861 status quo of serfdom was questioned after informational shocks in the 1850s created uncertainty. The second subsection shows how a middle ground mixing serfdom and liberal reform from 1861-1905 represented a learning trap in response to this uncertainty. The third subsection argues that, after another large shock in 1905, the government experimented with liberal reform, illustrating a third case of the model. The final subsection addresses two prominent alternative explanations for the learning trap — elite power and state capacity.

4.1 First Equilibrium: Serfdom Pre-1861

Prior to 1861, most Russian peasants provided services or payments to the state or a noble in exchange for use of arable land (peasants working noble land were "serfs").[12] Officials oversaw a "commune" of households in charge of repartitioning land, solving disputes, and managing resources.[13] Many peasants possessed limited control over production decisions, had few rights, and could be tied to their land in the event of sale[14] (Vasudevan (1988, 209), Nafziger (2010, 382), Nafziger (2016, 775-776) Dennison (2014, 252-257)).

While there were doubts about serfdom's viability, they had not "challenged the basic structure of local institutions" prior to the 1850s (Starr, 2015, 52). A minister of internal affairs contrasted Russia's "uninterrupted tranquility" under serfdom with the "[i]nternal discord and revolts" plaguing Europe's experiments with liberalism (Polunov et al., 2015, 87); moreover, the Russian countryside was not beset by an endemic economic crisis (Moon, 2014, 19-22). Russia hence represented an extremal policy (serfdom) under lower policy uncertainty.

Crimean War and Policy Uncertainty Russia's defeat in the 1853 Crimean War crystallized a need to modernize Russia's economy; many argued Russia lost the war due to economic backwardness. I interpret the defeat as a shock that generated uncertainty about

[12]Peasants rarely owned land, but both before and after 1861 owned small pieces of private property. There was heterogeneity in this arrangement; Nafziger (2012) shows how peasants on State-owned lands possessed some more individual property rights than serfs.

[13]Additionally, a failure of one commune member to pay taxes could result in punishment for others; or, the purchase of new land for the commune would make everyone collectively liable for paying mortgages on this land.

[14]For example, if noble A sold land to noble B, the serfs working that land would then be under the tutelage of noble B.

whether serfdom was inferior to Western European economic reform. Debates about economic reform circulated, for the first time, in public discourse (Starr, 2015, 57). Many educated Russians believed private economic activity was superior to the commune (Starr, 2015, 53-63), while other nobles favored a Manchesterian arrangement where individual peasants would rent and work nobles' land "on the basis of free competition" (Khristoforov, 2009, 65). Views that communal property was superior to private property due to "Russian exceptionalism" were derided by some economists (Kingston-Mann, 1991, 23-24).

Others maintained that serfdom and the commune were superior. Slavophilic ideologues romanticized the peasant as a self-sufficient "bearer of his own culture" (Khristoforov, 2009, 62) and argued the commune should be left untouched. Some believed private property could lead to a countryside marred with disorder, requiring seigneurial intervention to prevent exploitation of commons (Pravilova, 2014). Others contended that peasants were not aware of issues with the pre-1861 system (Field, 1976, 53), even arguing that peasants preferred communal land (Kingston-Mann, 1991, 34).

Regime's Preferences The Russian regime, headed by Tsar Alexander II, likely preferred a liberal path to modernization. A widening military gap with its neighbors was attributed to the stifled emergence of Russian industry due to the incentive structure of serfdom (Zenkovsky (1961, 284), Emmons (1966, 48-50)), as well as an inability to draw up army rseerves from the peasantry (Moon, 2014, 53-54). Abolishing serfdom could increase worker mobility and productivity, generating industrial growth that could fuel military prowess (Pereira, 1980, 108). Crucically, the Tsar believed the government should "begin eliminating serfdom from above [rather] than to wait until it begins to eliminate itself from below" (Pereira, 1980, 104).

A core motivation for preferring a liberal system was its fiscal efficiency and capacity to accelerate the development of state institutions (Dennison, 2023, 642). Under serfdom, tax collection was inefficiently outsourced to nobles who siphoned surpluses (Dennison, 2020, 2023). Liberal reform would weaken the nobility's bite, increasing revenue for the State. (Moon, 2014, 33).[15] If liberal reform could *substantially* increase surpluses, the government would be in a better place, even if it had to compensate elites for dispossessing their land. Finally, reformers hoped to use fiscal revenues to increase state capacity and fund the expansion of corporate forms and military ventures (Dennison, 2023, 642). The "Additional Results" online appendix develops a microfoundation for these fiscal preferences.

[15] Section 4.4 more carefully examines the nobility's impact on the reform process.

4.2 Second Equilibrium: Learning Trap, 1861-1905

This section shows how the post-1861 emancipation reforms and the codification of the commune can be viewed as a learning trap. I show how the government feared liberal reform might backfire, and how an intermediate policy between serfdom and liberal reform was conceived as a response to this fear, distorting economic incentives and halting learning about policy efficacy. Indeed, the relative intermediacy of the reform can even be thought of through the lens of unanticipated information revelation in section 3.3: the Crimean shock favored liberal reform, but since Russia began in a situation where serfdom was relatively accepted, the Tsar's policy did not radically swing in the direction of liberal reform.

Fears of Information Revelation Despite ideological or fiscal preferences for granting peasants economic freedom, Alexander II's government worried liberal policies could fail to *sufficiently* grow the Russian economy, viewing "communal land tenure and mobility restrictions [as] necessary precautions" in the face of this uncertainty (Dennison, 2014, 62). External validity arguments that Westernization was inappropriate for Russia accompanied fears of "the rise of an industrial proletariat, ruthless competition, unemployment, and other evils of industrial capitalism" (Polunov et al., 2015, 61).

Policy uncertainty was linked to fiscal uncertainty. Successful reform would fuel state capacity and the economy. In one state of the world, economic freedom would not only increase peasant income — via increases in agricultural productivity and the takeup of manufacturing jobs — but also ease revenue collection, generating fiscal growth even after compensating the nobility. However, the government was concerned that, in a non-ideal state of the world, alienation of rural land could limit its ability to collect the revenue necessary to compensate the nobility for expropriation, stressing state capacity (Dennison, 2023, 644). One of Alexander II's committees concluded that "the Order of the State might be shaken" by an emancipation that did *not* tie peasants to land (Field, 1976, 75), suggesting that the commune's institutionalization was seen as a solution to the ambiguous fiscal effects of reform.

Terms of Emancipation The Tsar's emancipation reforms beginning in 1861 ended serfdom and freed peasants from the coercive status-quo of seigneur-peasant agriculture. The reform impacted serfs working private lands more than state peasants, whose institutional constraints arguably changed less post-1861 (Nafziger, 2012). While broadening civil rights (Dennison, 2014), the reform did *not* remove the peasant commune and counterintuitively used it to constrain household labor mobility and capacity to acquire private property.

Nobles were compensated by the State for relinquishing land that was then redistributed between members of peasant communes. Ex-serfs made mortgage-like redemption payments to the State over 49 years to finance nobles' compensation. After payments were made, ex-serfs would privately own their allotted land but were otherwise collectively responsible for redemption and tax payments. Households could not legally sell, lease, transfer, alienate, or use land as collateral *until* redemption payments were complete, restricting freedom to abandon agriculture (Dower and Markevich (2019, 234), Nafziger (2012, 5-6), Nafziger (2016, 776), (Dennison, 2014, 259)).

Intermediate Policy The 1861 reforms mixed communal serfdom with liberal institutions. An external seigneur no longer regulated economic decisions. Peasants could privatize communal land by paying off redemption obligations (Nafziger, 2010, 383). Peasants owned houses, gardens, and rights to output (Kingston-Mann, 1991, 35). Communes (and sometimes households) even engaged in market land transactions with formal credit backing (Nafziger, 2010, 384). However, the persistence of the commune and its codification by the State clashed with individual economic incentives. Land allotments, repartitions, and redemption obligations were enacted communally, sometimes against households' wills. Peasants owned scattered strips of land, at best allowing for diversification, at worst increasing travel times and disincentivizing technological experimentation, slowing agricultural productivity growth (Williams (2013, 52-53, 65-66), Dower and Markevich (2019, 243)).

Mobility Restrictions The commune restricted labor mobility and supply by controlling the issuance of documents to pursue outside work (Nafziger, 2012, 7). Because households were communally responsible for redemption payments, those with larger obligations were incentivized to enforce mobility restrictions and free-ride on others. Nafziger (2010) notes: "Income-maximizing peasant households may have been unable to freely allocate resources between agricultural production and non-agricultural activities if they were subject to collective decision-making. . . By potentially inhibiting land transfers. . . and restricting the allocation of labor outside the village, the communal system may have introduced wedges between the shadow and market values for these factors of production" (384). Notably, these frictions stifled the flow of labor to cities, dampening the emergence of Russia's industrial sector and contradicting one of the original motives for reform.[16]

Sustained Uncertainty Over the decades, elites' and the State's inability to conclude whether the commune would have been better than individualized agriculture suggests that

[16] Indeed, Nafziger (2010) argues that weaker communal structures around Moscow allowed this area to develop a stronger industrial sector than other parts of Russia.

the post-1861 middle-ground sustained policy uncertainty. Despite the commune's apparent discouragement of technological adoption, Kingston-Mann (1991) documents that when innovation did occur, it *spread* more quickly in districts with greater communal intensity. While many in the government still believed in the pursuit of a liberal, Western path, one sources cites a government official,"an enthusiastic critic of the commune" (Kingston-Mann, 1991, 49), who conceded that the commune's ability to provide mutual aid during emergencies was a benefit that could not be realized under private property. Decades later, official contended that limited cases where peasants were permitted to sell land caused exploitation of peasant decision making (Wcislo, 2014, 89), casting doubt about the appropriateness of markets that would accompany liberal reform. Crucially, the *general equilibrium* effects of reform likely differed from the partial equilibrium experience of individual villages toying with private incentives. Regional heterogeneity in the enforcement of communal institutions allowed some households to observe the power of private incentives or accumulate wealth by participating in land rental markets or full-time non-agricultural labor (Nafziger, 2010, 385). Nevertheless, the contemporary remarks above suggest high-level uncertainty about whether liberal reform was surely better, even decades after 1861.

Implications Many contemporaries felt the Tsar's reforms did not go far enough. Finkel et al. (2015), for example, show a marked increase in peasant disturbances after 1861 due to the tepidness (relative to expectation) of the reform. More puzzlingly, in practice,"[f]ew surpluses could be wrung from peasants who were already obligated to repay the nobility for their freedom and their land and who continued to face obstacles to the most efficient allocation of their resources" (Dennison, 2023, 644-5). These realities make the choice to maintain the commune odd; if liberal reform could certainly result in large economic gains, the State would have been better off fiscally than in a middle ground where both peasant surplus and the ability to pay the nobility seemed constrained. But under policy uncertainty, fears that information revelation could strain the government's coffers and reputation were it to backfire give an informational explanation for this middle-ground.

From early on, the Tsar's government explicitly believed that a system that ending many of serfdom's abuses but still tying peasants to agricultural land could mitigate informational threats posed by more liberal reform. By sustaining uncertainty, a learning trap minimizes information revelation that could lead to overthrow. Not only did the Tsar's government hold power for decades, but except for the spike in disturbances following the Emancipation reform, reported peasant disturbances remained low in the following years (Finkel et al., 2015, 1000). In this sense, an "intermediate" policy, despite being inefficient in the interim, was successful in tiding threats of future overthrow from either radical liberal reform or allowing serfdom to remain.

However, this arrangement may have led to persistent and negative side-effects. Beyond stifling Russia's industrial sector and potentially slowing agricultural adoption in the medium term, Markevich and Zhuravskaya (2018) go as far as to argue that the allocation of property rights to communes instead of households dampened Russian economic growth into the 20th century. [17]

4.3 Third Equilibrium: Stolypin Reforms, Post-1905

Few radical modifications to the post-1861 arrangement occurred over the ensuing decades until the reforms headed by Prime Minister Pyotr Stolypin beginning in 1906, which represent experiments with private property.

These reforms were motivated by two forces which represent a large shock to beliefs in favor of liberal reform. The first was the culmination of investigations into the affairs of the peasantry led by Sergei Witte, the Russian Minister of Finance, whose findings — along with independent recommendations of the Ministry of Internal Affairs — favored the privatization of property and "individualized farming" (Yaney, 1964, 278). The second was the Russian Revolution of 1905 — peasant uprisings that led to the establishment of (limited) consitutional monarchy (Polunov et al., 2015, 223) and the suspension of households' redemption payments, making the reformation of social order — changing the relationship of the peasantry to the land — necessary to complement these new political institutions; indeed, the reforms existed "within the context of other legislative and organizational measures that Nicholas II's government introduced" (Pallot, 1999, 3). The confluence of these forces meant that "a sufficient number of officials in Nicholas II's government shared this conviction [that individualized farming should be introduced into the countryside] for it to find its way into official state policy" (Pallot, 1999, 4). This shock can also be viewed through the lens of unanticipated information shocks, as in section 3.3: a large shock in favor of liberal reform finally motivated a sharp shift in experimentation.

Stolypin's 1906 Reform gave households the choice between holding property in private or communal tenure. Households could leave the commune while maintaining allotments or an equivalent parcel of land, overseen by a state-mediated bargaining process. The dissolution of the commune's power was accompanied by a reduction in the enforcement of mobility restrictions. Stolypin indeed believed that peasants who found themselves landless would

[17]Some historians point to redemption payments as the first order inefficiency impeding livelihood in post-1861 ex-serf communities. Interestingly, Gerschenkron (1962) suggests that the commune *worsened* the bite of redemption payments. Households with particularly large debt obligations would have an incentive to free-ride on other members of the commune, affecting commune enforcement of labor mobility and production incentives which undoubtedly affected households' very ability to make payments.

find industrial work in the cities. From 1906 until the outbreak of the First World War, over one third of peasant households withdrew from the commune (Polunov et al., 2015, 231-32).

Stolypin's reforms were seen as experiments. The means of achieving land enclosure and alignment of rural social order with politically liberal institutions evolved as officials learned about peasants' willingness to take up the reform and grappled with administration (Yaney, 1964, 275-276, 283). The culmination of the process would, ideally, result in economic growth for Russia, and "[o]nce enlightened through education and example to the possibilities of organizing their farms differently, it was thought that the majority of Russia's peasants would greet the reform enthusiastically" (Pallot, 1999, 5), creating "an independent, prosperous husbandman, a stable citizen of the land" (1).

The Stolypin reforms represent a third and final case of the model, where experimentation with liberal reform occurs only after a shock sufficiently favors liberal reform. While the eventual course of the reform was cut short by the First World War, many argued that these reforms put Russia on the path towards economic development in imitation of the Western European model. Dower and Markevich (2019) argues that, indeed, the reform had a positive effect on the productivity of land-use in its short lifespan.

4.4 Alternative Interpretations for the Learning Trap

While Stolypin's experiments and pre-1862 serfdom can easily be seen as extreme, informative policies, the literature discusses two main *non-informational* arguments motivating the Tsar's 1861 reforms. First, this mixture may have balanced the desires of elites with those of the state. Second, Russia's lack of state capacity may have prevented more radical reform. In this section, I show that while these factors undoubtedly affected the shape of the emancipation reforms, the threat of informational revelation still affected the post-1861 policy outcome.

Elite Power The landed nobility's commitment to serfdom constrained previous attempts to abolish serfdom. Could the post-1861 reforms have simply been an outcome of bargaining between a liberal State and pro-serfdom nobles?

The commune may have been logistically preferable for dealing with civil and property disputes (Khristoforov and Gilley, 2016, 12). Labor allocation and rental payments may have been easier for landlords to organize when serfs were held collectively responsible (Dennison, 2014, 257). Landlords often disapproved of arrangements forcing them to give land to peasants, and extensively bargained over the size of government compensation during the years of the reform (Moon, 2014, chaps. 7-8). Nevertheless, the constraint binding the government *to the commune* was far weaker in the 1850s than decades prior.

Moreover, nobles, especially in the 1850s, were ideologically diverse, and may not have overwhelmingly wanted to keep the commune. In fact, many nobles' preferences aligned with the liberal aspirations of the Tsar. Most educated Russians believed that "the sanctity of private property was the basis of the political and social order" (Field, 1976, 58). Some liberals supported abolishing serfdom and endowing peasants with land. Even some conservative nobles wanted peasants to rent and work noble land in a competitive market. Many critics of the post-1861 system — at times more critical of serfdom than the government's own reformers — contrasted Russia's system of land tenure with the more desirable British case, "where large-scale land tenure and local self-government controlled by the aristocracy allegedly guaranteed economic prosperity and political stability" (Khristoforov, 2009, 58, 63-65). Khristoforov (2022) goes as far as to argue that some form of "private property individualism" was in the interest of both landlords and the government.[18]

Finally, a sufficient bloc of the nobility may not have wanted to challenge the Tsar; they "did not want political change... They wanted the tsar's favor... At each stage in the development of the government's program, the nobility swallowed its objections or stated them obliquely" (Field, 1976, 362).

State Capacity State capacity undoubtedly constrained the Tsar's reform ambitions. Emancipating and endowing peasants with property, compensating nobles, setting up a cadastre, and securing property rights was expensive. Depressed stock prices and banking crises in the 1850s drained coffers (Dennison, 2020, 197), meaning the government had to look to the gentry for help. Landlords then played a large role in commune politics after emancipation (Dennison, 2020); outsourcing dispute resolution and tax collection to the commune was logistically easier than protecting individual household rights (Khristoforov (2022, S163), Dennison (2011)).

Puzzles still arise from ascribing the intermediacy of the reform solely to weak administration. First, before and after the 1850s, the government could have relaxed the commune's power on State lands, where elites had a much weaker hold. Instead, the institutional experience of the State peasantry faced few changes before and after 1861, and outcomes of ex-Serf and State peasants converged by 1900 (Nafziger, 2012). Second, the State made few attempts to demarcate property. Officials wanted to tie peasants to the land, knew private property was easier to abandon through sale or mortgage, and in fact refused to separate the tax liabilities of some peasants who paid off their redemption obligations early (Khristoforov, 2007, 30). Dennison (2006) shows how the post-1861 commune sustained a pre-serfdom informal economy of social networks and patronage, even though the state could have slowly

[18] One reason was that a liberal system could oppose socialism.

introduced "universal enforcement of contracts and property rights" (89).

The state's nebulous attitude towards demarcation emanated in part from an informational desire to limit market activity. The results of a project attempting to construct a land cadastre prior to 1861 under Kiselev were never clearly revealed or publicly accessible, likely because the results may not have favored rationalization of property (Khristoforov and Gilley, 2016, 9, 11). Fourth, while the Tsar's committees "recognized [the commune] as a temporary fact" (Khristoforov, 2007, 30), Alexander II never took further steps towards dissolution of the commune, and his son infamously strengthened its enforcement powers [19] (Wortman, 1989, 21). Finally, the reforms were intended to *increase* fiscal revenue and expand state capacity. If liberal reform were truly beneficial for peasant income, maintaining the commune seems surprising. However, the government's belief that liberal reform could have shrunk agricultural output and caused land alienation suggests that uncertainty played a crucial role in the post-1861 mixture.

The power of a critical mass of the nobility, idiosyncratic financial issues, and administrative costs partially bound Russia's 1861 reforms, but the first was relaxed by the late 1850s, the second does not explain the uniformity of the State peasant experience, and the third does not explain why Russia experienced little liberal movement for decades, especially if the reform had fiscal objective. These limitations can be addressed by viewing the 1861 reforms as being in part driven by the risk of revealing information about policy efficacy.

5 Anticipated Information Revelation

This section extends the baseline model to show that, under threats of *anticipated* external information revelation, experimentation increases from both a direct channel (exogenous information) and an indirect channel (the leader herself is more likely to prefer experimentation, increasing learning). I show how this insight allows us to understand the interplay of policy contagion — where information about the effects of policies spreads from one polity to another — and political fragmentation. Specifically, I interpret 19th century Europe as a region where the probability of external information revelation was high and China as a region where it was low, and argue that the model can explain why European nations experimented with indirect rule and experienced high political turnover in the process, while China did not. Proofs of this section, as well as an additional extension showing the robustness of the model to multiple states of the world, are provided in the "Additional Results" part of

[19] For example, "[a]dditional legislation in 1893 explicitly forbade all sales of allotment land to non-peasants" and required a two-thirds majority of communal assembly votes to redeem or release landholdings and increased the portion of communal assembly votes (to two- thirds) required for a household to obtain release from their share or to redeem their land individually (Nafziger, 2010, 393).

the online appendix.

5.1 Information Revelation

Information revelation may be *anticipated* when leaders are consciously aware of external sources of inference. Officials in regions plagued by earthquakes or hurricanes may foresee that these disasters will eventually occur and reveal information about disaster relief policies. Another application is policy contagion, where a group of countries with similar institutions can learn from their policy experiences. Suppose that all these countries are stuck in a learning trap except for one, which commits to implement informative policies. Neighboring countries use these informative policies as an additional source of policy inference, meaning their leaders now face an additional threat of information revelation from external sources.

I model anticipated information shocks as follows. Let $\eta \in (0,1)$ be small. With probability η at the end of every period, the true state of the world is exogenously revealed. An unconstrained version of the leader's problem, anticipating the exogenous revelation of information, can be written as:

$$\max_{x\in[0,1]} \frac{2\sigma(1-\delta)u_\ell(x)+\delta((1-\eta)\Delta(x)+2\sigma\eta)\Psi(q)}{2\sigma(1-\delta)+\delta((1-\eta)\Delta(x)+2\sigma\eta)}.$$

Note that when $x = x_{LT}$, the leader's value is now

$$\frac{(1-\delta)u_{LT}+\delta\eta\Psi(q)}{(1-\delta)+\delta\eta}$$

Implementing x_{LT} is no longer a "safe option" for the leader that shuts down learning. Anticipating that she can no longer *fully* control information, she may take extreme actions to enjoy more flow utility. This intuition is reflected in the following proposition.

Proposition 5. *Suppose η increases. Then, $\underline{q}$ decreases and $\overline{q}$ increases. Fixing x', the leader plays the learning trap for fewer values of q.*

Anticipated information revelation improves policy learning compared to the unanticipated or baseline model. While there is mechanically more information revelation due to the realization of shocks, because $\underline{q}$ decreases, the leader's commits to experimental policies for a larger set of values, meaning learning also increases through an *endogenous* channel. As learning increases, q is more likely to move to 0. This means that fragmented regions plagued by policy contagion will see more governments conceding to their people — e.g. movements from absolute to constitutional monarchy — or, insofar as lower x is also a shorthand for increased external threats of overthrow, leaders' terms will be shorter, corresponding to more frequent political turnover.

5.2 Political Fragmentation in Europe and China

The model's prediction — politically fragmented regions of the world exhibit larger policy variation, more policy experimentation, and more frequent political turnover — is consistent with stylized historical facts differentiating Western Europe and China. It explains how the spread of constitutional monarchies and representative democracies in fragmented Western Europe contrasted with the consistent stability of consolidated Imperial China. Because systems of indirect rule were not widely experimented with by Eurasian states until the 1700s, we view our "uncertain policy" as the optimal degree of democracy or autocracy. The policy space corresponds to direct rule (e.g. monarchy) on one end and indirect rule (e.g. republican democracy) on the other end. Constitutional monarchies can be thought of as situations where autocrats begrudgingly cede power to the people but remain in office.

European countries around 1800 had a high η; Europe was populated by dozens of sovereign states — on average 85 between the years 1000 and 1799 (Dincecco and Wang, 2018)) — and thus possessed rich sources of policy information. England's "Glorious Revolution" at the end of the 17th century marked Europe's first constitutional monarchy (Kurian, 2011) and one of the first sources of data on indirect rule. After the French Revolution, rulers began to actively worry about citizens' desires for indirect rule (Evans and Von Strandmann, 2002, 10) as demands for freedom of speech, press, and disposal of private property became more prominent (Sperber, 2005, 66-67). By the first half of the 19th century, both Britain (constitutional monarchy) and America (representative democracy) were providing information on the efficacy of indirect rule. As these informational pressures strengthened, by 1850, most European states shifted to some form of government partially controlled by an elected legislature (Tilly, 1989). Some governments peacefully ceded power, while others had power wrested from them as part of the "Age of Revolutions." A threat of external information revelation forced governments to experiment with indirect rule, increased learning via convergence to indirect rule, and led to political upheaval.

In contrast, China was ruled by a single, unified State uninterrupted for a millennium. No states shared China's scale or political structure, suggesting it had low η in 1800. China possessed relatively integrated economic institutions that allowed its central government to oversee trade and a tightly-managed bureaucracy composed of strong elite networks (Rosenthal and Wong, 2011). The state's Confucian ideology centered political continuity and stability, contrasting with the "disruptive progress" of Western Europe. The entrenched power of the emperor coupled with a shared ideology of "stability" among the state and elites arguably prevented the emergence of ideologies promoting indirect rule, which was only strengthened by the lack of external information (Mokyr, 2016, ch. 16). China's high (although stagnant) economic standards in the late 1700s evidence a degree of policy *moderation* that was not seen in Western Europe (Pomeranz, 2021). Rare threats of overthrow came from regional elites and seemingly strengthened — rather than eroded — the collective power of the emperor and elites (Dincecco and Wang, 2018).

The "high η" environment of Western Europe in the 18th and 19th centuries was accompanied by greater policy variation, concessions to liberal demands, and political transitions, in contrast to the consolidated "low η" environment of China, which experienced little variation and minimal turnover. This prediction is consistent with Mokyr (2016), who argues that "there were repressive and reactionary regimes galore in Europe, but the interstate competitiveness constrained their ability to enforce a specific orthodoxy" (317). This prediction also complements a literature arguing that Europe's "political fragmentation" drove the Great Divergence by fostering competition over technology (Diamond, 2005), intellectual networks (Mokyr, 2016), and institutional innovation (Cox, 2017)

6 Conclusion

This paper studies a repeated environment where leaders set policy in the face of uncertainty. Agents learn about the effects of policies as they are implemented. Hence leaders' policies not only generate payoffs but also information which affects what policies they can implement in the future. Leaders facing threats of deposition are hence constrained by a threat of information revelation, generating an equilibrium nonmonotonicity in policy experimentation. When policies' effects are uncertain, leaders pursue spatially intermediate "learning trap policies" — shutting down information revelation when it would be most useful. They pursue extreme policies — keeping constituents indifferent to overthrow — only at extremal beliefs, when they are relatively certain about what the optimal policy is. Because learning traps deliver moderate payoffs, forever implementing a learning trap and revealing no information delivers a sure moderate payoff. Experimentation is preferred to these moderate payoffs only when beliefs that leaders' preferred policies favor their constituents are high. At intermediate beliefs, information is less likely to favor leaders in expectations; uncertainty about the optimal policy means they are able to shut down information revelation without being overthrown. At low beliefs, leaders would prefer to shut down information, but constituents are relatively certain about what the opitmal policy is and threaten overthrow if leaders do not implement it.

I use three equilbria of the model to study land reform in Imperial Russia. I argue that, when the Russian government was unsure about whether liberal economic institutions were preferable to serfdom, they implemented a "learning trap" — codifying the peasant commune and mixing elements of communal and liberal institutions. Even though the Russian government had ideological and fiscal preferences for liberal economic reform, they were uncertain about whether these policies would work in Russia due to cultural and institutional differences with Western Europe. However, by pursuing this intermediate policy, they distorted labor and production decisions, restricting economic, fiscal, and industrial growth but simultaneously obfuscating information about policy efficacy that minimized an informational threat of overthrow. I contrast this "learning trap" with Russia's adoption of pure serfdom prior to 1861 — in a setting where there was less policy uncertainty — and the liberal Stolypin experiments with private property after 1906 — representing experimentation with the government's preferred policy after a favorable informational shock.

The model's findings are robust to allowing for exogenous information arrival. Unanticipated information generates minimal policy variation unless a large amount of information favors the leader. Anticipated revelation generates greater policy variation since leaders cannot fully control information revelation. By interpreting exogenous information as the result of policy experimentation in an externally valid country, the latter insight can explain why Europe experienced a great deal of political upheaval and convergence to systems of indirect rule in the 19th century, while China did not, complementing "fragmentation" hypotheses of the Great Divergence.

Because learning traps arise at intermediate beliefs, learning about the optimal policy is shut down precisely when it would be most useful. This finding suggests that leaders who either lack term limits or prioritize keeping ideologically similar leaders in office will seldom experiment with potentially beneficial policies unless they are absolutely sure they will work in their favor. Future work will probe mechanisms that encourage efficient policy experimentation, as well as the impacts of polarization or alternative tenure systems, in such repeated settings.

References

Acemoglu, Daron and James A Robinson (2000). Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective. *The Quarterly Journal of Economics 115*(4), 1167–1199.

Acemoglu, Daron and James A Robinson (2001). A Theory of Political Transitions. *American Economic Review 91*(4), 938–963.

Acemoglu, Daron and James A Robinson (2008). Persistence of Power, Elites, and Institutions. *American Economic Review 98*(1), 267–93.

Aghion, Philippe and Matthew O Jackson (2016). Inducing Leaders to Take Risky Decisions: Dismissal, Tenure, and Term Limits. *American Economic Journal: Microeconomics 8*(3), 1–38.

Ashworth, Scott, Ethan Bueno De Mesquita, and Amanda Friedenberg (2017). Accountability and Information in Elections. *American Economic Journal: Microeconomics 9*(2), 95–138.

Banks, Jeffrey S and Rangarajan K Sundaram (1998). Optimal Retention in Agency Problems. *Journal of Economic Theory 82*(2), 293–323.

Besley, Timothy and Anne Case (1995). Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits. *The Quarterly Journal of Economics 110*(3), 769–798.

Boix, Carles and Milan W Svolik (2013). The Foundations of Limited Authoritarian Government: Institutions, Commitment, and Power-sharing in Dictatorships. *the Journal of Politics 75*(2), 300–316.

Buera, Francisco J, Alexander Monge-Naranjo, and Giorgio E Primiceri (2011). Learning the Wealth of Nations. *Econometrica 79*(1), 1–45.

Callander, Steven (2011). Searching for Good Policies. *American Political Science Review 105*(4), 643–662.

Chen, Yuyu and David Y Yang (2019). Historical Traumas and the Roots of Political Distrust: Political Inference from the Great Chinese Famine.

Cox, Gary W (2017). Political Institutions, Economic Liberty, and the Great Divergence. *The Journal of Economic History 77*(3), 724–755.

De Mesquita, Bruce Bueno and Alastair Smith (2010). Leader Survival, Revolutions, and the Nature of Government Finance. *American Journal of Political Science 54*(4), 936–950.

de Mesquita, Ethan Bueno and Wioletta Dziuda (2024). Partisan Traps.

Dennison, Tracy (2006). Did Serfdom Matter? Russian Rural Society, 1750–1860. *Historical Research 79*(203), 74–89.

Dennison, Tracy (2011). *The Institutional Framework of Russian Serfdom.* Cambridge University Press.

Dennison, Tracy (2014). *The Institutional Context of Serfdom in England and Russia* (NED - New edition ed.)., pp. 249–268. Boydell & Brewer.

Dennison, Tracy (2020). Overcoming Institutional Inertia: Serfdom, the State and Agrarian Reform in Prussia and Russia. In *A History of the European Restorations*, Volume 2, pp. 188–201. Bloomsbury Academic.

Dennison, Tracy (2023). Weak State, Strong Commune: Rural Authority in Imperial Russia. *Rivista Storica Italiana 135*(2), 595–622.

Dewan, Torun and Rafael Hortala-Vallve (2019). Electoral Competition, Control and Learning. *British Journal of Political Science 49*(3), 923–939.

Diamond, Jared M (2005). *Guns, Germs, and Steel: The Fates of Human Societies*. Norton.

Dincecco, Mark and Yuhua Wang (2018). Violent Conflict and Political Development over the Long Run: China versus Europe. *Annual Review of Political Science 21*, 341–358.

Dower, Paul Castañeda, Evgeny Finkel, Scott Gehlbach, and Steven Nafziger (2018). Collective Action and Representation in Autocracies: Evidence from Russia's Great Reforms. *American Political Science Review 112*(1), 125–147.

Dower, Paul C and Andrei Markevich (2019). The Stolypin Reform and Agricultural Productivity in Late Imperial Russia. *European Review of Economic History 23*(3), 241–267.

Dur, Robert AJ (2001). Why Do Policy Makers Stick to Inefficient Decisions? *Public Choice 107*(3-4), 221–234.

Eliaz, Kfir and Ran Spiegler (2020). A Model of Competing Narratives. *American Economic Review 110*(12), 3786–3816.

Ely, Jeffrey C and Juuso Välimäki (2003). Bad Reputation. *The Quarterly Journal of Economics 118*(3), 785–814.

Emmons, Terence LeRoy (1966). *The Russian Landed Gentry and the Peasant Emancipation of 1861.* University of California, Berkeley.

Evans, Robert John Weston and Hartmut Pogge Von Strandmann (2002). *The Revolutions in Europe, 1848-1849: From Reform to Reaction.* Oxford University Press on Demand.

Field, Daniel (1976). *The End of Serfdom: Nobility and Bureaucracy in Russia, 1865-1861.* Harvard University Press.

Finkel, Evgeny, Scott Gehlbach, and Tricia D Olsen (2015). Does Reform Prevent Rebellion? Evidence from Russia's Emancipation of the Serfs. *Comparative Political Studies 48*(8), 984–1019.

Fu, Qiang and Ming Li (2014). Reputation-Concerned Policy Makers and Institutional Status Quo Bias. *Journal of Public Economics 110*, 15–25.

Gerschenkron, Alexander (1962). *Economic Backwardness in Historical Perspective (1962).* The Belknap Press of Harvard University Press.

Holmström, Bengt (1999). Managerial Incentive Problems: A Dynamic Perspective. *The Review of Economic Studies 66*(1), 169–182.

Izzo, Federica (2024). With Friends Like These, Who Needs Enemies? *The Journal of Politics 86*(3), 835–849.

Izzo, Federica, Gregory J Martin, and Steven Callander (2021). Ideological Competition. *SocArXiv. February 19.*

Kartik, Navin, Francesco Squintani, Katrin Tinn, et al. (2015). Information Revelation and Pandering in Elections. *Columbia University, New York 36*, 8.

Khristoforov, IA (2007). The Fate of Reform: The Russian Peasantry in Government Policy and Public Opinion from the Late 1860s to the Early 1880s. *Russian Studies in History 46*(1), 24–42.

Khristoforov, Igor' (2009). Nineteenth-Century Russian Conservatism: Problems and Contradictions. *Russian Studies in History 48*(2), 56–77.

Khristoforov, Igor (2022). From Speransky to Stolypin: Peasant Reform and the Problem of Land Management. *Herald of the Russian Academy of Sciences 92*(3), S161–S173.

Khristoforov, Igor and Christopher Gilley (2016). Blurred Lines. *Cahiers du Monde Russe 57*(1), 31–54.

Kingston-Mann, Esther (1991). *Peasant Communes and Economic Innovation: A Preliminary Inquiry*, pp. 1–51. Princeton University Press.

Kurian, George (2011). Constitutional Monarchy. *The Encyclopedia of Political Science*, 316–317.

Levy, Gilat and Ronny Razin (2021a). A Maximum Likelihood Approach to Combining Forecasts. *Theoretical Economics 16*(1), 49–71.

Levy, Gilat and Ronny Razin (2021b). Short-Term Political Memory and the Inevitability of Polarisation. *Unpublished*.

Li, Yuan, Mario Gilli, et al. (2014). Accountability in Autocracies: The Role of Revolution Threat. Technical report, Stockholm School of Economics, Stockholm China Economic Research Institute.

Majumdar, Sumon and Sharun W Mukand (2004). Policy Gambles. *American Economic Review 94*(4), 1207–1222.

Manso, Gustavo (2011). Motivating Innovation. *The Journal of Finance 66*(5), 1823–1860.

Markevich, Andrei and Ekaterina Zhuravskaya (2018). The Economic Effects of the Abolition of Serfdom: Evidence from the Russian Empire. *American Economic Review 108*(4-5), 1074–1117.

Mokyr, Joel (2016). *A Culture of Growth*. Princeton University Press.

Montiel Olea, José Luis, Pietro Ortoleva, Mallesh M Pai, and Andrea Prat (2022, 04). Competing Models. *The Quarterly Journal of Economics*. qjac015.

Moon, David (2014). *The Abolition of Serfdom in Russia: 1762-1907*. Routledge.

Mukand, Sharun W and Dani Rodrik (2005). In Search of the Holy Grail: Policy Convergence, Experimentation, and Economic Performance. *American Economic Review 95*(1), 374–383.

Nafziger, Steven (2010). Peasant Communes and Factor Markets in Late Nineteenth-Century Russia. *Explorations in Economic History 47*(4), 381–402.

Nafziger, Steven (2012). Serfdom, Emancipation, and Off-Farm Labour Mobility in Tsarist Russia. *Economic History of Developing Regions 27*(1), 1–37.

Nafziger, Steven (2016). Communal Property Rights and Land Redistributions in Late Tsarist Russia. *The Economic History Review 69*(3), 773–800.

Pallot, Judith (1999). *Land Reform in Russia, 1906-1917: Peasant Responses to Stolypin's Project of Rural Transformation.* Clarendon press.

Pereira, NGO (1980). Alexander II and the Decision to Emancipate the Russian Serfs, 1855-61. *Canadian Slavonic Papers 22*(1), 99–115.

Polunov, Aleksandr, Thomas C Owen, and Larissa Georgievna Zakharova (2015). *Russia in the Nineteenth Century: Autocracy, Reform, and Social Change, 1814-1914: Autocracy, Reform, and Social Change, 1814-1914.* Routledge.

Pomeranz, Kenneth (2021). *The Great Divergence.* Princeton University Press.

Prat, Andrea (2005). The Wrong Kind of Transparency. *American Economic Review 95*(3), 862–877.

Pravilova, Ekaterina (2014). *A Public Empire: Property and the Quest for the Common Good in Imperial Russia.* Princeton University Press.

Prendergast, Canice and Lars Stole (1996). Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning. *Journal of Political Economy 104*(6), 1105–1134.

Roland, Gerard (2002). The Political Economy of Transition. *Journal of Economic Perspectives 16*(1), 29–50.

Rosenthal, Jean-Laurent and R Bin Wong (2011). *Before and Beyond Divergence: The Politics of Economic Change in China and Europe.* Harvard University Press.

Schwartzstein, Joshua and Adi Sunderam (2021). Using Models to Persuade. *American Economic Review 111*(1), 276–323.

Sperber, Jonathan (2005). *The Pre-Revolutionary Political Universe* (2 ed.)., pp. 56?108. New Approaches to European History. Cambridge University Press.

Spiegler, Ran (2016). Bayesian Networks and Boundedly Rational Expectations. *The Quarterly Journal of Economics 131*(3), 1243–1290.

Starr, Frederick S (2015). *Decentralization and Self-Government in Russia, 1830-1870*. Princeton University Press.

Stokey, Nancy L, Robert E Lucas, and Edward C Prescott (1989). *Recursive Methods in Economic Dynamics*. Harvard University Press.

Tilly, Charles (1989). State and Counterrevolution in France. *Social Research*, 71–97.

Tomasi, Arduino (2023). The Stakes and Informativeness Trade-Off: Electoral Incentives to Implement Programmatic Transfers. *The Journal of Politics 85*(2), 654–666.

Vasudevan, Hari S (1988). Peasant Land and Peasant Society in Late Imperial Russia. *The Historical Journal 31*(1), 207–222.

Wang, Shaoda and David Y Yang (2021). Policy Experimentation in China: The Political Economy of Policy Learning. Technical report, National Bureau of Economic Research.

Wcislo, Francis William (2014). *Reforming Rural Russia*. Princeton University Press.

Williams, Stephen F (2013). *Liberal Reform in an Illiberal Regime: The Creation of Private Property in Russia, 1906-1915*. Number 545. Hoover Institution Press.

Wortman, Richard (1989). Property Rights, Populism, and Russian Political Culture. *Civil Rights in Imperial Russia*, 13–32.

Yaney, George L (1964). The Concept of the Stolypin Land Reform. *Slavic Review 23*(2), 275–293.

Zenkovsky, Serge A (1961). The Emancipation of the Serfs in Retrospect. *Russian Review*, 280–293.

Proofs of Main Results

Proof of Solution to People's Problem

Proposition 3. *$x_p(q,x')$ is the solution to the following Bellman equation:*

$$W(q,x') = \max_{x\in[0,1]} \frac{2\sigma(1-\delta)\mathbb{E}[y|x] + \delta\Delta(x)\Phi(q,x)}{2\sigma(1-\delta)+\delta\Delta(x)} - k|x-x'|$$

where $\Phi(q,x) = qW(1,x) + (1-q)W(0,x)$. For intermediate values of c, there exists a nonempty open set $LT \subseteq [0,1]\times X$ with $LT \ni (1/2, x_{LT})$ such that for all $(q,x') \in LT$, $W(q,x') \le y_{LT} + c \le W(0,x'), W(1,x')$.

Lemma 1. *Suppose x^* is a solution to the functional equation defining $W(q,x')$. Then, x^* is also a solution to the functional equation defining $W(q,x^*(q,x'))$.*

Proof. This proof utilizes the linearity of adjustment costs. Suppose by contradiction that x^{**} solves the functional equation associated with $W(q,x^*)$ but not $W(q,x')$. Since both x^* and x^{**} are always feasible, this implies

$$(1-\delta)\mathbb{E}[y|x^{**}] - k|x^{**}-x^*| + \frac{\delta\Delta(x^{**})}{2\sigma}(qW(1,x^{**}) + (1-q)W(0,x^{**})) + \delta(1-\frac{\Delta(x^{**})}{2\sigma}W(q,x^{**})$$
$$> (1-\delta)\mathbb{E}[y|x^*] + \frac{\delta\Delta(x^*)}{2\sigma}(qW(1,x^*) + (1-q)W(0,x^*)) + \delta(1-\frac{\Delta(x^*)}{2\sigma})W(q,x^*)$$

which also implies

$$(1-\delta)\mathbb{E}[y|x^{**}] - k|x^{**}-x^*| - k|x^*-x'| + \frac{\delta\Delta(x^{**})}{2\sigma}(qW(1,x^{**}) + (1-q)W(0,x^{**}))$$
$$+\delta(1-\frac{\Delta(x^{**})}{2\sigma})W(q,x^{**})$$
$$> (1-\delta)\mathbb{E}[y|x^*] - k|x^*-x'| + \frac{\delta\Delta(x^*)}{2\sigma}(qW(1,x^*) + (1-q)W(0,x^*)) + \delta(1-\frac{\Delta(x^*)}{2\sigma})W(q,x^*)$$

However, since $k|x^{**}-x'| \le k|x^{**}-x^*| + k|x^*-x'|$, this implies

$$(1-\delta)\mathbb{E}[y|x^{**}] - k|x^{**}-x'| + \frac{\delta\Delta(x^{**})}{2\sigma}(qW(1,x^{**}) + (1-q)W(0,x^{**})) + \delta(1-\frac{\Delta(x^{**})}{2\sigma})W(q,x^{**})$$
$$\ge (1-\delta)\mathbb{E}[y|x^{**}] - k|x^{**}-x^*| - k|x^*-x'| + \frac{\delta\Delta(x^{**})}{2\sigma}(qW(1,x^{**}) + (1-q)W(0,x^{**}))$$
$$+\delta(1-\frac{\Delta(x^{**})}{2\sigma})W(q,x^{**})$$
$$> (1-\delta)\mathbb{E}[y|x^*] - k|x^*-x'| + \frac{\delta\Delta(x^*)}{2\sigma}(qW(1,x^*) + (1-q)W(0,x^*)) + \delta(1-\frac{\Delta(x^*)}{2\sigma})W(q,x^*)$$

contradicting that x^* solved the functional equation associated with $W(q, x')$. □

Lemma 2. *There exist* $0 \le \underline{x} < x_{LT} < \overline{x} < \tilde{x} \le \overline{\overline{x}} \le 1$ *such that*

$$W(1, x') = \begin{cases} f_g(\overline{x}) - k(\overline{x} - x') & x' \notin (\overline{x}, \overline{\overline{x}}) \\ f_g(x') & x' \in (\overline{x}, \overline{\overline{x}}) \end{cases}$$

$$W(0, x') = \begin{cases} f_b(\underline{x}) - k(x' - \underline{x}) & x' \ge \underline{x} \\ f_b(x') & x' < \underline{x} \end{cases}$$

Proof. Suppose $q = 1$ and $x' \le \tilde{x}$. $W(1, x') = \max_{x \in [0,1]}(1 - \delta) f_g(x) + \delta W(1, x) - k|x - x'|$ The previous lemma allows us to simplify this to $\max_{x \in [0,1]} f_g(x) - k|x - x'|$, which is solved by $\overline{x} = \tilde{x} - k/2$. A reverse argument shows $\overline{\overline{x}} = \tilde{x} + k/2$. That these points are to the right of x_{LT} follows from $f_g'(x_{LT}) > k$. □

Lemma 3. *Suppose* x^* *is a solution to the functional equation defining* $W(q, x')$ *and* $q \in (0, 1)$ *Then:*

$$W(q, x^*) = \frac{2\sigma(1 - \delta)\mathbb{E}[y|x^*] + \delta\Delta(x^*)\Phi(q, x^*)}{2\sigma(1 - \delta) + \delta\Delta(x^*)}$$

where $\Phi(q, x) = qW(1, x) + (1 - q)W(0, x)$.

Moreover, $W(q, x') = \max_x \frac{2\sigma(1-\delta)\mathbb{E}[y|x] + \delta\Delta(x^*)\Phi(q,x)}{2\sigma(1-\delta)+\delta\Delta(x)} - k|x - x'|$.

Proof. Since x^* solves the functional equation associated with $W(q, x')$, by the previous lemmata it also solves the one associated with $W(q, x^*)$. Hence:

$$\begin{aligned} W(q, x^*) &= (1 - \delta)\mathbb{E}[y|x^*] + \frac{\delta\Delta(x^*)}{2\sigma}\Phi(q, x^*) + \delta(1 - \frac{\Delta(x^*)}{2\sigma})W(q, x^*) \\ &= \frac{(1 - \delta)\mathbb{E}[y|x^*] + \frac{\delta\Delta(x^*)}{2\sigma}\Phi(q, x^*)}{1 - \delta(1 - \frac{\Delta(x^*)}{2\sigma})} = \frac{2\sigma(1 - \delta)\mathbb{E}[y|x^*] + \delta\Delta(x^*)\Phi(q, x^*)}{2\sigma(1 - \delta) + \delta\Delta(x^*)} \end{aligned}$$

$W(q,x')$ can then be written as:

$$\begin{aligned}W(q,x') &= \max_{x^*}(1-\delta)\mathbb{E}[y|x^*] + \frac{\delta\Delta(x^*)}{2\sigma}\Phi(q,x) \\ &\quad + \delta(1-\frac{\Delta(x^*)}{2\sigma})\frac{2\sigma(1-\delta)\mathbb{E}[y|x^*]+\delta\Delta(x^*)\Phi(q,x)}{2\sigma(1-\delta)+\delta\Delta(x)} - k|x^*-x'| \\ &= \max_x \mathbb{E}[y|x]\big((1-\delta)+\delta(1-\frac{\Delta(x)}{2\sigma})\frac{2\sigma(1-\delta)}{2\sigma(1-\delta)+\delta\Delta(x)}\big) \\ &+ \delta\Phi(q,x)\big(\frac{\Delta(x)}{2\sigma}+(1-\frac{\Delta(x)}{2\sigma})\frac{\delta\Delta(x)}{2\sigma(1-\delta)+\delta\Delta(x)}\big) - k|x-x'| \\ &= \max_x \mathbb{E}[y|x]\big((1-\delta)\big(1+\frac{2\sigma\delta-\delta\Delta(x)}{2\sigma(1-\delta)+\delta\Delta(x)}\big)\big) \\ &+ \delta\frac{\Delta(x)}{2\sigma}\Phi(q,x)\big(1+\frac{2\sigma\delta-\delta\Delta(x)}{2\sigma(1-\delta)+\delta\Delta(x)}\big) - k|x-x'| \\ &= \max_x \frac{2\sigma(1-\delta)\mathbb{E}[y|x]+\delta\Delta(x)\Phi(q,x)}{2\sigma(1-\delta)+\delta\Delta(x)} - k|x-x'|\end{aligned}$$

□

Lemma 4. *Suppose $x' \in [\underline{x}, \overline{x}]$. Then $x_p(q,x') \in [\underline{x}, \overline{x}]$.*

Proof. Note that our problem can be written as:

$$\begin{aligned}W(q,x') = \max_{x\in[0,1]} &\frac{2\sigma(1-\delta)}{2\sigma(1-\delta)+\delta\Delta(x)}\big(\mathbb{E}[y|x]-k|x-x'|\big) \\ &+ \frac{\delta\Delta(x)}{2\sigma(1-\delta)+\delta\Delta(x)}\big(\Phi(q,x)-k|x-x'|\big)\end{aligned}$$

First, since $\frac{\partial^2}{\partial q \partial x}\mathbb{E}[y|x] = f_g'(x) - f_b'(x) > 0$, since $\arg\max f_g(x) - k|x-x'| = \overline{x}$, and since $\arg\max f_b(x) - k|x-x'| = \underline{x}$, we have: $\arg\max_{\hat{x}\in[0,1]} \mathbb{E}[y|\hat{x}] - k|\hat{x}-x'| \in [\underline{x}, \overline{x}]$. All $x < \underline{x}$ are dominated by $\underline{x}$ and all $x > \overline{x}$ are dominated by $\overline{x}$.

Next, I claim $\arg\max_{\tilde{x}\in[0,1]} W(q,\tilde{x}) - k|\tilde{x}-x'| \in [\underline{x}, \overline{x}]$. Suppose $x < \underline{x}$, meaning $f_b(x)$ is concave and $f_b'(x) > -k$ for $x < \underline{x}$. Then, the derivative with respect to x here is: $qk + (1-q)f_b'(x) + k = (1-q)(f_b'(x)+k) > 0$, meaning $\underline{x}$ dominates all options to its left. Similarly, for $x \in (\overline{x}, \overline{\overline{x}})$, $|f'(x)|$ is necessarily $< k$ and the derivative is $qf'(x) - (1-q)k - k < 0$. Finally, for $x \geq \overline{\overline{x}}$, the derivative is $-k-k < 0$ meaning $\overline{\overline{x}}$ dominates in this range, but since the left and right derivatives at $\overline{\overline{x}}$ are then both negative, we have that $\overline{x}$ dominates all options to its right.

Finally, suppose by contradiction that $[\underline{x}, \overline{x}] \not\ni \hat{x}$, where $\hat{x}$ is the maximizer associated with $W(q,x')$. If $\hat{x} < \underline{x}$, we know that $\mathbb{E}[y|\hat{x}] - k|\hat{x}-x'| < \mathbb{E}[y|\underline{x}] - k|\underline{x}-x'|$ and $\Phi(q,\hat{x}) - k|\hat{x}-x'| < \Phi(q,\underline{x}) - k|\underline{x}-x'|$ Any convex combination of $\mathbb{E}[y|\underline{x}] - k|\underline{x}-x'|$ and $\Phi(q,\underline{x}) - k|\underline{x}-x'|$ is strictly larger than any convex combination of $\mathbb{E}[y|\hat{x}] - k|\hat{x} - x'|$ and $\Phi(q, \hat{x}) - k|\hat{x} - x'|$, meaning playing $\overline{x}$ dominates $\hat{x}$ in the original program. A symmetric argument shows that playing any $\hat{x} > \overline{x}$ is always strictly worse than $\overline{x}$. □

Lemma 5. *Under Assumptions 1 and 2', there exists $c > 0$ such that: $f_g(\overline{x}) - k(\overline{x} - x_{LT}) \geq y_{LT}, f_b(\underline{x}) - k(x_{LT} - \underline{x}) + c > y_{LT}$ and $W(1/2, x_{LT}) < y_{LT} + c$. Hence, the set $LT = \{(q, x') : W(q, x') < y_{LT} - c\}$ has nonempty interior.*

Proof. We constructing an upper bound for $W(1/2, x_{LT})$. First, I claim:

$$\frac{1}{2}\big(f_g(\overline{x}) - k(\overline{x} - x_{LT})\big) + \frac{1}{2}\big(f_b(\underline{x}) - k(x_{LT} - \underline{x})\big) - k(\overline{x} - x_{LT}) < y_{LT}$$
$$\frac{1}{2}\big(f_g(\overline{x}) - k(\overline{x} - x_{LT})\big) + \frac{1}{2}\big(f_b(\underline{x}) - k(x_{LT} - \underline{x})\big) - k(x_{LT} - \underline{x}) < y_{LT}.$$

Sufficient conditions for the first equation are: $f_g(\overline{x}) - 2k(\overline{x} - x_{LT}) < y_{LT}$ and $f_b(\underline{x}) - k(\overline{x} - \underline{x}) < y_{LT}$. The first inequality follows immediately from the fact that $f_g'(x_{LT}) < 2k$, contained in Assumption 2. Next, note that $x_{LT} = \frac{\beta_b - \beta_g + \tilde{x}^2}{2\tilde{x}}$. Then, using the fact that $\underline{x} = k/2$ and $\overline{x} = \tilde{x} - k/2$, $f_b(\overline{x}) - k(\overline{x} - \underline{x}) < y_{LT}$ if and only if:

$$\beta_b - \frac{k^2}{4} - k(\tilde{x} - k) < \beta_b - \left(\frac{\beta_b - \beta_g + \tilde{x}^2}{2\tilde{x}}\right)^2 \iff \frac{3}{4}k^2 - \tilde{x}k + \left(\frac{\tilde{x}^2}{4} + (\beta_b - \beta_g) + \frac{(\beta_b - \beta_g)^2}{4\tilde{x}^2}\right) < 0$$

This holds if $k \in \left[\frac{2\tilde{x} - \sqrt{\tilde{x}^2 - 3\tilde{x}^2/4 - 3(\beta_b - \beta_g) - 3(\beta_b - \beta_g)^2/\tilde{x}^2}}{3}, \frac{2\tilde{x} + \sqrt{\tilde{x}^2 - 3\tilde{x}^2/4 - 3(\beta_b - \beta_g) - 3(\beta_b - \beta_g)^2/\tilde{x}^2}}{3}\right]$ and $k \in \left[\frac{2\tilde{x} - \sqrt{\tilde{x}^2/4 - 3(\beta_b - \beta_g) - 3(\beta_b - \beta_g)^2/\tilde{x}^2}}{3}, \frac{2\tilde{x} + \sqrt{\tilde{x}^2/4 - 3(\beta_b - \beta_g) - 3(\beta_b - \beta_g)^2/\tilde{x}^2}}{3}\right]$. This is satisfied by Assumptions 2 and 2', which also subsume a similar condition with the second inequality:
$k \in \left[\frac{2\tilde{x} - \sqrt{\tilde{x}^2/4 + 3(\beta_b - \beta_g) - 3(\beta_b - \beta_g)^2/\tilde{x}^2}}{3}, \frac{2\tilde{x} + \sqrt{\tilde{x}^2/4 + 3(\beta_b - \beta_g) - 3(\beta_b - \beta_g)^2/\tilde{x}^2}}{3}\right]$.

We now construct our upper bound for $W(1/2, x_{LT})$. $W(1/2, x_{LT}) =$

$$\max_{x \in [\underline{x}, \overline{x}]} \frac{2\sigma(1-\delta)\mathbb{E}[y|x] + \delta\Delta(x)\Phi(1/2, x)}{2\sigma(1-\delta) + \delta\Delta(x)} - k|x - x_{LT}|$$
$$= \max_{x \in [\underline{x}, \overline{x}]} \frac{2\sigma(1-\delta)}{2\sigma(1-\delta) + \delta\Delta(x)}\big(\mathbb{E}[y|x] - k|x - x_{LT}|\big) + \frac{\delta\Delta(x)}{2\sigma(1-\delta) + \delta\Delta(x)}\big(\Phi(1/2, x) - k|x - x'|\big)$$
$$\leq \max_{x \in [\underline{x}, \overline{x}]} \frac{2\sigma(1-\delta)}{2\sigma(1-\delta) + \delta\Delta(x)}\big(\max_{\breve{x}} \mathbb{E}[y|\breve{x}] - k|\breve{x} - x_{LT}|\big) + \frac{\delta\Delta(x)}{2\sigma(1-\delta) + \delta\Delta(x)}\big(\Phi(1/2, x) - k|x - x'|\big)$$
$$= \max_{x \in [\underline{x}, \overline{x}]} \frac{2\sigma(1-\delta)}{2\sigma(1-\delta) + \delta\Delta(x)} y_{LT} + \frac{\delta\Delta(x)}{2\sigma(1-\delta) + \delta\Delta(x)}\big(\Phi(1/2, x) - k|x - x_{LT}|\big) \equiv \max_{x \in [\underline{x}, \overline{x}]} U(x)$$

$U(x)$ will be the objective associated with the upper bound. The third line holds since

$$\partial/\partial x\Big|_{x=x_{LT}} \mathbb{E}[y|x] = \frac{1}{2}2(\tilde{x} - x_{LT}) - \frac{1}{2}2x_{LT} = \tilde{x} - 2x_{LT} = \tilde{x} - 2\frac{\tilde{x}^2 + (\beta_b - \beta_g)}{2\tilde{x}} = \frac{\beta_b - \beta_g}{2\tilde{x}}$$

and by Assumption 2', $|\frac{\beta_b - \beta_g}{2\tilde{x}}||\frac{\beta_b - \beta_g}{\tilde{x}}| < |\frac{\beta_b - \beta_g}{\tilde{x} - k}| < k$.

Note that Φ does not depend on x at $q = 1/2$: $1/2(f_g(\overline{x}) - k(\overline{x} - x)) + 1/2(f_b(\underline{x}) - k(x - \underline{x})) = 1/2(f_g(\overline{x}) + f_b(\underline{x}) - k(\overline{x} - \underline{x}))$. Hence:

$$\begin{aligned}\Phi(1/2, x_{LT}) =& 1/2(f_g(\overline{x}) - k(\overline{x} - x_{LT})) + 1/2(f_b(\underline{x}) - k(x_{LT} - \underline{x})) > y_{LT} \\ >& 1/2(f_g(\overline{x}) - k(\overline{x} - x_{LT})) + 1/2(f_b(\underline{x}) - k(x_{LT} - \underline{x})) - k(\overline{x} - x_{LT}), \\ & 1/2(f_g(\overline{x} - k(\overline{x} - x_{LT})) + 1/2(f_b(\underline{x}) - k(x_{LT} - \underline{x})) - k(\underline{x} - x_{LT})\end{aligned}$$

Let $\hat{x}_1$ and $\hat{x}_2$ be defined by the following:

$$\begin{aligned} y_{LT} &= 1/2(f_g(\overline{x} - k(\overline{x} - x_{LT})) + 1/2(f_b(\underline{x}) - k(x_{LT} - \underline{x})) - k(\hat{x}_1 - x_{LT}) \qquad \hat{x}_1 \in (x_{LT}, \overline{x}) \\ y_{LT} &= 1/2(f_g(\overline{x} - k(\overline{x} - x_{LT})) + 1/2(f_b(\underline{x}) - k(x_{LT} - \underline{x})) - k(\hat{x}_2 - x_{LT}) \qquad \hat{x}_2 \in (\underline{x}, x_{LT}) \end{aligned}$$

which exist and are unique by continuity, strict monotonicity, and the Intermediate Value Theorem. Because we are taking convex combinations of y_{LT} and $\Phi - k|x - x_{LT}|$, where the weights depend on x as well, we will only want to place less weight on y_{LT} if $\Phi - k|x - x_{LT}| > y_{LT}$ (precisely on $[\hat{x}_2, x_{LT}]$ and $[x_{LT}, \hat{x}_1]$).

The derivative of $U(x)$ for $x > x_{LT}$ is:

$$\frac{2\sigma(1-\delta)\delta\Delta'(x)(\Phi - k(x - x_{LT}) - y_{LT}) - 2\sigma(1-\delta)\delta\Delta(x)k - \delta^2\Delta(x)^2 k}{(2\sigma(1-\delta) + \delta\Delta(x))^2}$$

A nearly identical derivative for $x < x_{LT}$ is omitted. At $x = x_{LT}$, the right derivative is $\frac{\delta\Delta'(x)(\Phi - y_{LT})}{2\sigma(1-\delta)} > 0$. At $\hat{x}_1$, when $\Phi - k(\hat{x}_1 - x_{LT}) - y_{LT} = 0$, the derivative is $-\frac{\delta\Delta(x)k}{2\sigma(1-\delta)+\delta\Delta(x)} < 0$. The derivative is ≥ 0 when:

$$\begin{aligned} 2\sigma(1-\delta)\delta\Delta'(x)(\Phi - k(x - x_{LT}) - y_{LT}) &\geq \delta\Delta(x)(2\sigma(1-\delta) + \delta\Delta(x))k \\ \Phi - k(x - x_{LT}) - y_{LT} &\geq \frac{\delta\Delta(x)(2\sigma(1-\delta) + \delta\Delta(x))k}{2\sigma(1-\delta)\delta\Delta'(x)} \\ \Phi - k(x - x_{LT}) &\geq y_{LT} + \frac{\Delta(x)(2\sigma(1-\delta) + \delta\Delta(x))}{2\sigma(1-\delta)\Delta'(x)}k \\ &= y_{LT} + k(x - x_{LT})\left(1 + \frac{\delta\Delta(x)}{2\sigma(1-\delta)}\right) \end{aligned}$$

The final line is strictly increasing in x with a zero at $x_{LT} < \hat{x}_1$. $\Phi - k(x - x_{LT})$ is strictly decreasing in x with a zero at $\hat{x}_1$. Hence there is a *unique* maximizer $\overline{x}^\star$ to this problem in $(x_{LT}, \hat{x}_1)$. Define c implicitly as $\Phi - k(\overline{x}^\star - x_{LT}) = y_{LT} + c$. Defining c using the derivative to the left of x_{LT} gives the same c, since Φ, $\Delta(x)$, and $k|x - x_{LT}|$ are symmetric about x_{LT}. Then:

$$W(1/2, x_{LT}) \leq \max_x U(x) = y_{LT} + \frac{\delta\Delta(\overline{x}^\star)}{2\sigma(1-\delta) + \delta\Delta(\overline{x}^\star)}\left(k(\overline{x}^\star - x_{LT})\left(1 + \frac{\delta\Delta(\overline{x}^\star)}{2\sigma(1-\delta)}\right)\right)$$
$$= y_{LT} + \frac{\delta\Delta(\overline{x}^\star)}{2\sigma(1-\delta) + \delta\Delta(\overline{x}^\star)} c < y_{LT} + c.$$

Since $x^\star > x_{LT}$, we also have $\Phi(1/2, x_{LT}) - k(x^\star - x_{LT}) = y_{LT} + c \implies \Phi(1/2, x_{LT}) > y_{LT} + c$. That c is not *so* large that the people are discouraged from overthrowing at $q = 0, 1$ follows from the fact that $\Phi(1/2, x_{LT}) = \frac{1}{2}W(1, x_{LT}) + \frac{1}{2}W(0, x_{LT})$ and that $|\beta_g - \beta_b|$ is small. Since $W(q, x')$ is continuous, the inequality $W(1/2, x_{LT}) < y_{LT} + c$ also holds in a *neighborhood* of $(1/2, x_{LT})$. This means that $LT : \{(q, x') : W(q, x') \leq y_{LT} + c\}$ has nonempty interior. □

Proof of Solution to Leader's Problem

Theorem 1. *Suppose $\sigma(1 - \delta^\ell)$ is small or $\underline{u}(x')$ is sufficiently low. Then, there exist thresholds $\underline{q} < \overline{q} \in (0, 1)$ such that the following hold:*

1. *If $q \leq \underline{q}$, then $x_\ell(q, x') = \underline{x}_{LT}(q, x')$ or $\overline{x}_{LT}(q, x')$. The leader plays the* closest policy *to the learning trap.*

2. *If $q \in [\underline{q}, \overline{q}]$, then $x_\ell(q, x') = \underline{x}_{LT}(q, x')$ or $\overline{x}(q, x')$. The leader plays either the lowest policy below the learning trap or her highest feasible policy.*

3. *If $q > \overline{q}$, = then $x_\ell(q, x') = \overline{x}(q, x')$ or $\underline{x}(q, x')$ (and it is necessary that $\Delta(\underline{x}(q, x')) > \Delta(\overline{x}(q, x'))$). The leader either plays her highest feasible policy or policies maximizing information revelation.*

Lemma 6. *Suppose that the leader can implement any $x \in [x_\ell(0, x'), x_\ell(1, x')]$. Then, there exist thresholds $\underline{q} < \overline{q}$ such that: for $q \leq \underline{q}$, the derivative of the objective in the maximization problem is positive for $x < x_{LT}$ and negative for $x > x_{LT}$; for $q \in [\underline{q}, \overline{q}]$, the derivative of the objective is positive for $x < x_{LT}$ and positive for $x > x_{LT}$; and for $q \geq \overline{q}$, the derivative of the objective is negative for $x < x_{LT}$ and positive for $x > x_{LT}$.*

Proof. Let $X_{unc} = [x_\ell(0, x'), x_\ell(1, x')]$. The leader's Bellman in the unconstrained case can be written as:

$$\tilde{V}(q) = \max_{x \in X_{unc}} (1-\delta)u_\ell(x) + \delta\frac{\Delta(x)}{2\sigma}(q\overline{u}(x') + (1-q)\underline{u}(x')) + \delta\left(1 - \frac{\Delta(x)}{2\sigma}\right)\tilde{V}(q)$$

Let $\Psi(q) = q\overline{u}(x') + (1-q)\underline{u}(x')$. Replicating an earlier argument implies

$$\tilde{V}(q) = \max_{x \in X_{unc}} \frac{2\sigma(1-\delta)u_\ell(x) + \delta\Delta(x)\Psi(q)}{2\sigma(1-\delta) + \delta\Delta(x)}.$$

The numerator of the derivative with respect to x for $x \neq x_{LT}$ is:

$$= 4\sigma^2(1-\delta)^2 u'_\ell(x) + 2\sigma\delta(1-\delta)(\Delta(x)u'_\ell(x) + \Delta'(x)\Psi(q) - \Delta'(x)u_\ell(x)).$$

Observe that $\Delta(x)$ is the absolute value of a linear function with zero at x_{LT}.

$$\Delta(x) = |\beta_g - (x-\tilde{x})^2 - (\beta_b - x^2)| = |\beta_g - \beta_b - \tilde{x}^2 + 2\tilde{x}x| \equiv \xi|x - x_{LT}|$$

This means that for $x > x_{LT}$, the derivative of the objective is negative if and only if

$$2\sigma(1-\delta)u'_\ell(x) \leq \delta\big(\Delta'(x)u_\ell(x) - \Delta(x)u'_\ell(x) - \Delta'(x)\Psi(q)\big)$$
$$2\sigma(1-\delta)u'_\ell(x) \leq \delta\xi\big(u_\ell(x) - u'_\ell(x)(x - x_{LT}) - \Psi(q)\big)$$

$u_\ell(x) - u'_\ell(x)(x - x_{LT})$ is equal to u_{LT} at $x = x_{LT}$ and is weakly increasing for $x > x_{LT}$.[20] By weak concavity, the LHS below is decreasing and the RHS increasing:

$$2\sigma(1-\delta)u'_\ell(x) \leq \delta\xi\big(u_\ell(x) - u'_\ell(x)(x - x_{LT}) - \Psi(q)\big) \tag{1}$$

If $\sigma(1-\delta)$ is small, since $\underline{u}(x') < u_{LT} \leq u_\ell(x) - u'_\ell(x)(x - x_{LT}) < \overline{u}(x')$, there is a threshold $\underline{q}$ such that the inequality satisfies strictly if and only if $q < \underline{q}$. Moreover, for any $\sigma(1-\delta)$, as $\underline{u}(x')$ grows more negative, $-\Psi(q)$ grows large, which can also allow the inequality to satisfy. A similar argument shows that for $x < x_{LT}$, the derivative is positive when:

$$2\sigma(1-\delta)u'_\ell(x) \geq \delta\xi\big(\Psi(q) - \big(u_\ell(x) - u'_\ell(x)(x - x_{LT})\big)\big) \tag{2}$$

When (1) holds with equality, its RHS is strictly positive, meaning the RHS of (2) is strictly negative, showing $\underline{q} < \overline{q}$. Hence: for $x > x_{LT}$, the derivative of the objective is strictly negative when $q < \underline{q}$ and strictly positive when $q > \underline{q}$, with equality at $q = \underline{q}$; for $x < x_{LT}$, the derivative

[20]To see this, note that the derivative of this expression with respect to x is $u'_\ell(x) - u''_\ell(x)(x - x_{LT}) - u'_\ell(x) = -u''_\ell(x)(x - x_{LT}) \geq 0$.

of the objective is strictly positive when $q < \overline{q}$ and strictly negative when $q > \overline{q}$, with equality at $q = \overline{q}$. This completes the lemma. □

Lemma 7. *Let $x_\ell(q, x')$ denote a policy for the leader such that $x_\ell(q, x_\ell(q, x')) = x_\ell(q, x')$. Then, the people overthrow if and only if $\tilde{W}(x_\ell, q, x') \geq W(q, x') - c$, where*

$$\tilde{W}(x_\ell, q, x') = \frac{2\sigma(1-\delta)\mathbb{E}[y|x_\ell(q,x')] + \delta\Delta(x_\ell(q,x'))\tilde{\Phi}(q, x_\ell(q,x'))}{2\sigma(1-\delta) + \delta\Delta(x)} - k|x_\ell(q,x') - x'|$$

where $\tilde{\Phi}(q, x) = q\big(f_g(x_\ell(1,x) - k|x_\ell(1,x) - x'|\big) + (1-q)\big(f_b(x_\ell(0,x) - k|x_\ell(0,x) - x|\big)$.

Proof. Let $\tilde{W}(x_\ell, q, x')$ be the value for the people of accepting for all time a policy $x_\ell(q, x')$ satisfying the hypothesis. Let

$$\begin{aligned}
\tilde{\Phi}(q,x) &= q\tilde{W}(x_\ell, 1, x_\ell(1, x_\ell(1,x'))) + (1-q)\tilde{W}(x_\ell, 1, x_\ell(0, x_\ell(0,x'))) \\
\tilde{W}(x_\ell, 1, x') &= \Big(\sum_{t=0}^{\infty}(1-\delta)f_g(x_\ell(1,x'))\Big) - k|x_\ell(1,x') - x'| = f_g(x_\ell(1,x') - k|x_\ell(1,x') - x'| \\
\tilde{W}(x_\ell, 0, x') &= \Big(\sum_{t=0}^{\infty}(1-\delta)f_b(x_\ell(0,x'))\Big) - k|x_\ell(0,x') - x'| = f_b(x_\ell(0,x') - k|x_\ell(0,x') - x'|.
\end{aligned}$$

Then:

$$\begin{aligned}
\tilde{W}(x_\ell, q, x_\ell(q,x')) &= (1-\delta)\mathbb{E}[y|x_\ell(q,x')] + \delta\frac{\Delta(x_\ell(q,x'))}{2\sigma}\Phi(q, x_\ell(q,x')) \\
&\quad + \delta\Big(1 - \frac{\Delta(x)}{2\sigma}\Big)\tilde{W}(q, x_\ell(q,x')) \\
&= \frac{2\sigma(1-\delta)\mathbb{E}[y|x_\ell(q,x')] + \delta\Delta(x_\ell(q,x'))\Phi(q, x_\ell(q,x'))}{2\sigma(1-\delta) + \delta\Delta(x)}
\end{aligned}$$

Replicating arguments from earlier lemmata gives the result:

$$\tilde{W}(x_\ell, q, x') = \frac{2\sigma(1-\delta)\mathbb{E}[y|x_\ell(q,x')] + \delta\Delta(x_\ell(q,x'))\Phi(q, x_\ell(q,x'))}{2\sigma(1-\delta) + \delta\Delta(x)} - k|x_\ell(q,x') - x'|$$

□

Lemma 8. $x_\ell(q, x') = x_\ell\big(q, x_\ell(q, x')\big)$*, meaning the previous lemma characterizes the equilibrium overthrow decision.*

Proof. If $\text{NR}(q, x')$ does not bind, the leader implements her preferred policy by the previous lemma, which satisfies the property. So, suppose $\text{NR}(q, x')$ binds and that by contradiction there exists x' such that $x_1 \equiv x_\ell(q, x') \neq x_\ell\big(x_\ell(q, x')\big) \equiv x_2$. Let $h(x) = (1-\delta)\mathbb{E}[y|x] +$

$\frac{\delta\Delta(x)}{2\sigma}\tilde{\Phi}(q,x)$. Then, we must have

$$\begin{aligned} W(q,x') - c &= h(x_1) - k|x_1 - x'| \\ &\quad + \delta\big(1 - \frac{\Delta(x_1)}{2\sigma}\big)\big(h(x_2) + \delta\big(1 - \frac{\Delta(x_2)}{2\sigma}\big)\tilde{W}_{x_\ell}(x_2) - k|x_2 - x_1|\big) \\ W(q,x_1) - c &= h(x_2) + \delta\big(1 - \frac{\Delta(x_2)}{2\sigma}\big)\tilde{W}_{x_\ell}(x_2) - k|x_2 - x_1| \end{aligned}$$

Let $x^* = x_p(q,x')$. Since $\text{NR}(q,x')$ is binding, it must be that:

$$\begin{aligned} \frac{2\sigma(1-\delta)\mathbb{E}[y|x^*] + \delta\Delta(x^*)\Phi(q,x^*)}{2\sigma(1-\delta) + \delta\Delta(x^*)} - k|x^* - x'| - c > \\ h(x_2) + \delta\big(1 - \frac{\Delta(x_2)}{2\sigma}\big)\tilde{W}_{x_\ell}(x_2) - k|x_2 - x'| \end{aligned}$$

where the RHS is equal to $W(q,x')$. Since either $x^* < x_1 < x_2$ or $x^* > x_1 > x_2$ by a previous lemma (policy is attracted in a single direction; implementing x^* and then x_2 is dominated by a policy between x^* and x_2), as $x' \to x_1$, both sides increase/decrease by the same amount or the right side increases by more, meaning:

$$\frac{2\sigma(1-\delta)\mathbb{E}[y|x^*] + \delta\Delta(x^*)\Phi(q,x^*)}{2\sigma(1-\delta) + \delta\Delta(x^*)} - k|x^* - x_1| > h(x_2) + \delta\big(1 - \frac{\Delta(x_2)}{2\sigma}\big)\tilde{W}_{x_\ell}(x_2) - k|x_2 - x_1|$$

But then:

$$\begin{aligned} W(q,x_1) &= \max_{x\in[0,1]} \frac{2\sigma(1-\delta)\mathbb{E}[y|x] + \delta\Delta(x)\Phi(q,x)}{2\sigma(1-\delta) + \delta\Delta(x)} - k|x - x_1| \\ &\geq \frac{2\sigma(1-\delta)\mathbb{E}[y|x^*] + \delta\Delta(x^*)\Phi(q,x^*)}{2\sigma(1-\delta) + \delta\Delta(x^*)} - k|x^* - x_1| \end{aligned}$$

a contradiction.

□

Proof. We use the lemmata above to prove the original theorem; the solution can be written as:

$$V(q,x') = \max_{x\in\text{NR}(q,x')} \frac{2\sigma(1-\delta)u)\ell(x) + \delta\Delta(x)\Psi(q)}{2\sigma(1-\delta) + \delta\Delta(x)}$$

where, replicating the previous lemma's argument, we have $\Psi(q) = q\overline{u}(x') - (1-q)\underline{u}(x')$. If $q \leq \underline{q}$, the derivative for $x > x_{LT}$ is negative and the derivative for $x < x_{LT}$ is positive, meaning the leader is attracted to the learning trap. Hence, $x_\ell(q,x') = \underline{x}_{LT}(q,x')$ or $\overline{x}_{LT}(q,x')$. If $q \in [\underline{q},\overline{q}]$, the solution is either $\underline{x}_{LT}(q,x')$ or some $x \geq x_{LT}$. If $q \geq \overline{q}$. the derivative for $x > x_{LT}$

is positive and the derivative for $x < x_{LT}$ is negative. Hence the optimal solution is either $\overline{x}(q, x')$ or $\underline{x}(q, x')$. $\square$

Additional Results

Preference Microfoundation: Land Reform

Overview This subsection provides a microfoundation for the preferences of the "people" and "leader" with an eye to the fiscal setting of Russian land reform. Peasants produce output, which is extracted by the leader via taxation that is siphoned away by elites. We represent the people's utility as the sum of post-tax surplus for the gentry and peasantry. The main mechanism driving the leader's preference for disempowering the gentry — beyond effects on aggregate output by changing peasant production incentives — is that disenfranchising the gentry (high x in the main model) allows them to siphon less tax revenue.

We allow the preference of the leader, in this case, to depend on the state of the world, and relax the assumption of strict monotonicity; however, this does not radically alter the proofs in the previous section.

Setup Consider an economy made up of peasants, gentry, and leader. The union of the peasants and gentry compose the "people." Each period, peasants produce surplus $i(x) \geq 0$, depending on $x \in [0,1]$ and the state of the world (pro or anti-leader). $x = 0$ represents a system of overseen serf labor. $x = 1$ represents a system where peasant households are given private property and labor mobility. x_{LT} represents an arrangement where the commune remains and the peasantry have some economic freedom.

The leader collects revenue from the peasantry in a process intermediated by the gentry by setting a tax $t \in [0,1]$ on $i(x)$. Assume there is some maximal-level of taxation $\bar{t} < 1$ that the peasantry will permit (reflecting e.g. subsistence or overthrow threats) so that the leader sets $t = \bar{t}$ in equilibrium.

Since x reflects not only peasants' economic incentives but also gentry involvement, suppose that when the government taxes at t, the gentry siphons away a share $s(x) \in [0,1]$ of revenue. We assume s is strictly decreasing for $x < x_{LT}$ and constant for $x \geq x_{LT}$. For $x < x_{LT}$, the leader gradually expropriates gentry land, meaning they can siphon less-and-less. For $x \geq x_{LT}$, the leader only has to worry about compensating the gentry for their land (so $s(x)$ is constant but is likely > 0).

Writing y as Post-Tax Surplus Let $u_p(x)$ denote the sum of both peasants' and gentry utility:

$$u_p(x) = (1 - \bar{t} + \bar{t}s(x))i(x)$$

Suppose that for each state of the world, $i(x)$ =:

$$\frac{f_g(x)}{1-\bar{t}+\bar{t}s(x)} \qquad \text{Pro-Leader}$$
$$\frac{f_b(x)}{1-\bar{t}+\bar{t}s(x)} \qquad \text{Anti-Leader}$$

Assume that $f_g(x), f_b(x) \geq 0$ on $[0,1]$ since $i(x)$ represents output. With this interpretation, y_t in each stage of the main game represent the *post-tax* output left to the "people" (gentry plus peasantry); substituting these expressions for $i(x)$ back into $u_p(x)$ gives $u_p(x) = f_g(x)$ in the pro-leader state and $f_b(x)$ in the anti-leader state of the world. Note that since $f_g(x_{LT}) = f_b(x_{LT})$, $i(x_{LT})$ is the same in both states of the world; hence, both aggregate output and post-tax output are the same in each state of the world, and inference made with either quantity yields similar results. Next, since $1-\bar{t}+\bar{t}s(x)$ is weakly decreasing in x and positive, $i(x)$ is strictly increasing in x in the pro-leader state of the world, just as $f_g(x)$. The derivative with respect to x in the anti-leader state is

$$\frac{f_b'(x)}{1-\bar{t}+\bar{t}s(x)} - \bar{t}s'(x)\frac{f_b(x)}{\left(1-\bar{t}+\bar{t}s(x)\right)^2}$$

which is strictly decreasing for $x > x_{LT}$ and for some region of $x < x_{LT}$ as well. This means that *aggregate* output may not be higher under full-serfdom, but that departing from x_{LT} results in some gains, and that *post-tax surplus* for the sum of the elite and peasant surpluses is higher in this state of the world.

Establishing Leaders' Preferences The leader's revenue is given by $\bar{t}(1-s(x))i(x)$. Using the mapping above, utility $u_\ell(x)$ in each state of the world is

$$\bar{t}(1-s(x))\frac{f_g(x)}{1-\bar{t}+\bar{t}s(x)} \qquad \text{Pro-Leader}$$
$$\bar{t}(1-s(x))\frac{f_b(x)}{1-\bar{t}+\bar{t}s(x)} \qquad \text{Anti-Leader}$$

In the pro-leader state of the world, $u_\ell(x)$ is clearly strictly increasing in x. In the anti-leader state of the world, for the main results of the main model to go through, we simply require that

$$\bar{t}(1-s(x_\ell(0,x')))\frac{f_b(x_\ell(0,x'))}{1-\bar{t}+\bar{t}s(x_\ell(0,x'))} < \bar{t}(1-s(x_{LT}))\frac{y_{LT}}{1-\bar{t}+\bar{t}s(x_{LT})}$$

recalling that $x_\ell(0, x')$ is the solution to the leader's problem in the main model when the anti-leader state of the world. This holds, for example, if $s(x_\ell(0, x'))$ is large relative to $s(x_{LT})$, and would seem to hold in the Russian case, where the gentry have a strong hold in the peasantry during serfdom but have their property expropriated after the reforms (at x_{LT}).

These pieces exemplify precisely how a "divide-the-dollar" land-reform arrangement can serve as a foundation for the preferences of the leader in the model and how these preferences align with features of the 1861 Russian reforms.

External Information Revelation

Proposition 5. *Suppose η increases. Then, $\underline{q}$ decreases and $\overline{q}$ increases. Fixing x', the leader plays the learning trap for fewer values of q.*

Proof. The objective of an unconstrained problem here here can be written as:

$$\max_{x\in[0,1]} \frac{2\sigma(1-\delta)u_\ell(x) + \delta((1-\eta)\Delta(x) + 2\sigma\eta)\Psi(q)}{2\sigma(1-\delta) + \delta((1-\eta)\Delta(x) + 2\sigma\eta)}$$

The numerator of the derivative can be expressed as:

$$\begin{aligned}
=&\big(2\sigma(1-\delta) + \delta(1-\eta)\Delta(x) + 2\sigma\delta\eta\big)\big(2\sigma(1-\delta)u_\ell'(x) + \delta(1-\eta)\Delta'(x)\Psi(q)\big) \\
&- \delta(1-\eta)\Delta'(x)\big(2\sigma(1-\delta)u_\ell(x) + \delta\big((1-\eta)\Delta(x) + 2\sigma\eta\big)\Psi(q)\big) \\
=&4\sigma^2(1-\delta)^2u_\ell'(x) + 2\sigma\delta(1-\delta)(1-\eta)\big(\Delta(x)u_\ell'(x) + \Delta'(x)\Psi(q) - \Delta'(x)u_\ell(x))\big) \\
&+ 2\sigma\delta\eta\big(2\sigma(1-\delta)u_\ell'(x) + \delta(1-\eta)\Delta'(x)\Psi(q) - (1-\eta)\Delta'(x)\Psi(q)\big) \\
=&4\sigma^2(1-\delta)^2u_\ell'(x) + 2\sigma\delta(1-\delta)(1-\eta)\big(\Delta(x)u_\ell'(x) + \Delta'(x)\Psi(q) - \Delta'(x)u_\ell(x))\big) \\
&+ 4\sigma\delta(1-\delta)\eta u_\ell'(x)
\end{aligned}$$

For $x > x_{LT}$, the derivative is negative, replicating a previous lemma, if:

$$2\sigma(1-\delta)u_\ell'(x) + 2\sigma\delta\eta u_\ell'(x) \le \delta(1-\eta)\xi\big(u_\ell(x) - u_\ell'(x)(x - x_{LT}) - \Psi(q)\big)$$

where $\underline{q}$ is the unique point that solves this equation with equality. For each x, the LHS is increasing in η and the RHS is decreasing in η. This means that as η increases from 0, $\underline{q}$ shifts to the left. The equation defining $\overline{q}$ can be written as:

$$2\sigma(1-\delta)u_\ell'(x) + 2\sigma\delta\eta u_\ell'(x) = \delta(1-\eta)\xi\big(\Psi(\overline{q}) - \big(u_\ell(x) - u_\ell'(x)(x - x_{LT})\big)\big)$$

An increase in η likewise causes the LHS to increase and RHS to decrease, causing $\overline{q}$ to increase. $\square$

Three States of the World

This section shows robustness of this paper's fundamental insight on leaders' policies to three states of the world by establishing a version of Lemma 6. Analogues of Proposition 1 or Theorem 1 are difficult to work with because $\text{NR}(q, x')$ is hard to pin down.

The three states of the world are described as follows

$$\begin{aligned} f_g(x) \qquad & y_t = \beta_g - (x_t - \tilde{x}_g)^2 + \epsilon_t \\ f_m(x) \qquad & y_t = \beta_m - (x_t - \tilde{x}_m)^2 + \epsilon_t \\ f_b(x) \qquad & y_t = \beta_b - x_t^2 + \epsilon_t \end{aligned}$$

for $\tilde{x}_g > \tilde{x}_m \in (0, 1]$ and $\epsilon_t \sim \mathcal{U}[-\sigma, \sigma]$ in all cases, as earlier. We assume $f_i(x) - f_j(x)$ each single-cross at some x_{LT}^{ij} for all $i \neq j$. Examples are graphed below.

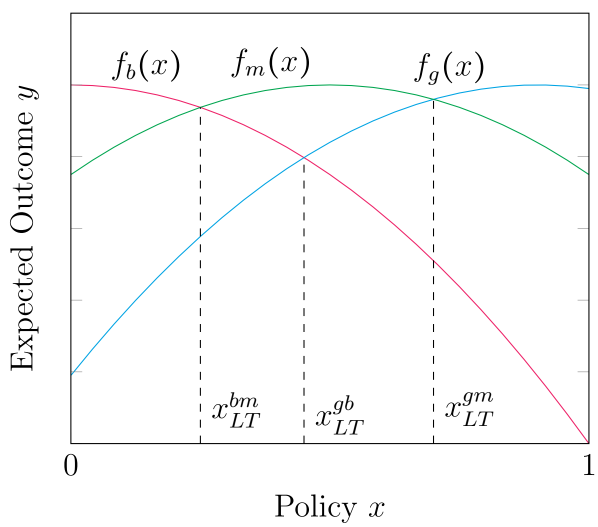

Figure 5: $f_g(x)$, $f_m(x)$, and $f_b(x)$ graphed

Our state variables are q_g (belief that g holds) and q_m (belief that m holds) and we fix $q_g, q_m > 0$ and $q_g + q_m < 1$ for the remainder of the analysis. Beginning at a vector of states (q_g, q_m), we can write the following expected values of the problem, as well as the probabilities with which they occur (which are all functions of x).

- V^{gmb}: initial value of the problem at (q_g, q_m).

- V^g: value of problem if g is true; $(q_g, q_m) \to (1, 0)$; occurs with prob. p_g. Solution is denoted x^g_ℓ.
- V^m: value of problem if m is true; $(q_g, q_m) \to (0, 1)$; occurs with prob. p_m Solution is denoted x^m_ℓ.
- V^b value of problem if b is true; $(q_g, q_m) \to (0, 0)$; occurs with prob. p_b. Solution is denoted x^b_ℓ.
- V^{gb} value of problem if only m is not true; $(q_g, q_m) \to (\frac{q_g}{1-q_m}, 0)$; occurs with prob. p_{gb}
- V^{gm} value of problem if only b is not true; $(q_g, q_m) \to (\frac{q_g}{q_g+q_m}, \frac{q_m}{q_g+q_m})$; occurs with prob. p_{gm}
- V^{mb} value of problem if only g is not true; $(q_g, q_m) \to (0, \frac{q_m}{1-q_g})$; occurs with prob. p_{mb}

Each of these probabilities can be computed explicitly in terms of the beliefs and x; for example, $p_g(x)$ is given by $q_g \cdot \frac{|f_g(x)-f_b(x)|}{2\sigma} \cdot \frac{|f_g(x)-f_m(x)|}{2\sigma}$.

We assume that $x^b_\ell < x^{gb}_{LT} < x^g_\ell$ and $x^b_\ell < x^{mb}_{LT} < x^m_\ell$. Fixing $1-q_m-q_b$, suppose the difference between $V^b = u_\ell(x^b_\ell)$ and $V^m = u_\ell(x^m_\ell)$ is sufficiently large. Applying Lemma 6 shows:

- $V^{gb} = x^{gb}_{LT}$: suppressing the risk of information revelation that could reveal b as true, the leader will choose a "pairwise" learning trap x^{gb}_{LT}
- $V^{mb} = x^{mb}_{LT}$: suppressing the risk of information revelation that could reveal b as true, the leader will choose a "pairwise" learning trap x^{mb}_{LT}

This gives us the following result:

Proposition 6. *Suppose the following: $q_g, q_m > 0$ and $q_g + q_m < 1$; the leader's policies are constrained if g, m, or b are revealed as true but can otherwise play any policy in $[x^b_\ell, x^g_\ell]$; and V^b is sufficiently small. Then, the leader either implements a learning trap policy x^{gb}_{LT} or x^{mb}_{LT}.*

Proof. The proof for this result is intuitive. Replicating earlier arguments, the leader's problem can be written as a convex combination of flow and continuation values:

$$V^{gmb} = \max_x \frac{(1-\delta)u_\ell(x) + \delta\big(p_g V^g + p_m V^m + p_b V^b + p_{gb} V^{gb} + p_{gm} V^{gm} + p_{mb} V^{mb}\big)}{(1-\delta) + \delta\big(p_g + p_m + p_b + p_{gb} + p_{gm} + p_{mb}\big)}$$

$$V^{gmb} = \max_x \frac{(1-\delta)u_\ell(x) + \delta\big(p_g V^g + p_m V^m + p_b V^b + p_{gb} x^{gb}_{LT} + p_{gm} V^{gm} + p_{mb} x^{mb}_{LT}\big)}{(1-\delta) + \delta\big(p_g + p_m + p_b + p_{gb} + p_{gm} + p_{mb}\big)}$$

Since we assume V^b is sufficiently small, the leader will do anything in her power to suppress revealing that b is true, i.e. will try its best to set $p_b = 0$. p_b is expressed as:

$$p_b = (1 - q_g - q_m)\frac{|f_g(x) - f_b(x)|}{2\sigma} \cdot \frac{|f_m(x) - f_b(x)|}{2\sigma}$$

With probability $1 - q_g - q_m$, b is true. Then, with probability $\frac{|f_g(x)-f_b(x)|}{2\sigma}$, state b is separable from state g and with probability $\frac{|f_g(x)-f_b(x)|}{2\sigma}$ it is separable from state m. Because the left and right derivatives of p_b are always nonzero, the only way to set p_b equal to zero is either to set $f_g(x) = f_b(x)$ by playing x_{LT}^{gb}, or to set $f_m(x) = f_b(x)$ by implementing x_{LT}^{mb}. The option that is chosen will depend on the parameters of the model (precise beliefs and values of x_{LT}^{gb} and V^m). □

The fundamental insight is that a leader, when faced with a threat of future overthrow, will *always* be attracted to a moderate policy that shuts down information revelation about b. More subtly, the terminal history of the model always reduces to one or two states of the world . Suppose the leader implements x_{LT}^{gb}, allowing information revelation only about whether $\{g, b\}$ or $\{m\}$ is true. In the former case, we arrive at the baseline two-state model. On the equilibrium path, uncertainty is resolved only when upon revelation of states to which the leader is not highly averse.

Model with Non-Uniform Distribution of ϵ_t

Finally, we return to the model with myopia and solve a version of the model assuming that $\epsilon_t \sim p$, where $p(\cdot)$ is continuous, has mean 0, is symmetric around 0, and single-peaked. Denote the distribution of posteriors q_t given a prior q_{t-1} and today's policy x_t as $\tau(q_t|q_{t-1}, x_t)$. Note that, if $|x_t' - x_{LT}| > |x_t - x_{LT}|$, for $x \leq \tilde{x}$, because the expected value of the posterior is the prior, $\tau(q_t|q_{t-1}, x_t')$ is a mean preserving spread of $\tau(q_t|q_{t-1}, x_t)$; that is, policies farther from x_{LT} reveal more information.

Next, note that in the model with myopia, $\text{NR}(q) = \text{NR}(q, x') = [\tilde{x}q - \sqrt{c}, \min\{\tilde{x}q + \sqrt{c}, 1\}]$ is convex in (q, x). If the people are not myopic and have a large preference for learning, this assumption may not necessarily hold, since the people simply prefer any extreme policy that reveals information.

The leader's problem can now be expressed as:

$$V(q) = \max_{x_\ell(q) \in \text{NR}(q)} \underbrace{(1 - \delta)u_\ell(x)}_{\text{Flow utility}} + \underbrace{\delta \int V(q')\tau(q'|q, x)}_{\text{Value of information}},$$

When the leader implements a policy x, this policy potentially generates information, in addition to flow utility. Note that, because the expected value of the posterior is the prior, the (continuation) value of information is simply $V(q)$ integrated over the distribution $\tau(q'|q,x)$. When $x = x_{LT}$, $\tau(q'|q,x)$ places mass 1 on the prior q; as x grows more extreme, the variance of posteriors increases. If V is concave in q, there is hence an aversion to information revelation.

The following proposition characterizes $x_\ell(q)$; the proof is identical for arbitrary $\text{NR}(q,x')$ as long as this set is convex.

Proposition 7. *There exists a threshold* $\underline{q} \in (0,1]$ *such that:*

1. *if* $q \le \underline{q}$, $x_\ell(q) = \overline{x}(q)$, *where* $\underline{q}$ *solves* $\tilde{x}\underline{q} + \sqrt{c} = x_{LT}$. *The leader plays the most extreme policy preventing overthrow.*

2. *If* $q \ge \underline{q}$, *the leader implements* $x_\ell(q) \in [\min\{x_{LT}, \underline{x}(q)\}, \overline{x}(q)]$. *Moreover, for* δ *high,* $x_\ell(q) < \overline{x}(q)$; *i.e., she moderates her actions to pursue the learning trap.*

Proof. Suppose q is such that $\overline{x}(q) \le 1$. Because $\text{NR}(q)$ is convex in (q,x), by Theorem 9.8 in Stokey et al. (1989), $V(q)$ is concave in q. In particular, V's concavity in the belief q suggests a general *aversion* to information revelation, and hence a tendency to pursue policies closer to x_{LT} whenever a leader values this information sufficiently, i.e. when she is patient.

Let $\underline{q} \in (0,1)$ solve $\tilde{x}\underline{q} + \sqrt{c} = x_{LT}$. Suppose $q \le \underline{q}$. Note that, on this range, an increase from x to x' has two effects. First, it increases flow utility $u_\ell(q)$. Second, because $q \le \underline{q}$, we must have $x_{LT} \ge x' > x$, so $|x - x_{LT}| > |x' - x_{LT}|$, meaning $\tau(q'|q,x')$ is a mean-preserving contraction of $\tau(q'|q,x)$. This means that, because $V(q)$ is concave in q, $\int V(q')\tau(q'|q,x') > \int V(q')\tau(q'|q,x)$, i.e. the value of information increases. Hence, the leader will pursue $\overline{x}(q)$, which is both the highest possible policy and the closest policy to the learning trap. This shows the first part of the proposition.

Next, suppose $q > \underline{q}$. By the previous argument, all feasible $x < x_{LT}$ are dominated by x_{LT}. Now, above x_{LT}, an increase from x to x' has two effects. First, it increases flow utility: $u_\ell(x') > u_\ell(x)$. Second, because V is concave, it *decreases* the value of continuation utility, i.e. $\int V(q')\tau(q'|q,x') < \int V(q')\tau(q'|q,x)$. For δ large, this second informational effect will tend to outweigh the former flow effect, suggesting that $\min\{\underline{x}(q), x_{LT}\} \le x_\ell(q) < \overline{x}(q)$. □