Dynamic Reform of Public Institutions: A Model of Motivated Agents and Collective Reputation*

Justin Valasek†

July 18, 2017

Abstract

State capacity is optimized when public institutions are staffed by individuals with public-service motivation. However, when motivated agents value the collective reputation of their place of employment, steady-state equilibria with both high and low aggregate motivation (reputation) in the mission-oriented sector exist. Reforming a low-motivation institution requires a non-monotonic wage path: since the effect of higher wages on motivation is negative for a high-reputation institution, but positive for a low-reputation institution, a transition to a high-reputation steady state requires an initial wage increase to crowd motivated workers in, followed by a wage decrease to crowd non-motivated workers out.

Keywords: Motivated Workers, Institutional Reform, Public Sector.
JEL Classification Codes: D23, D73, L32.

*Thanks to Rachel Kranton, Steffen Huck, Paul Heidhues, Rajshri Jayaraman, Sebastian Schweighofer-Kodritsch, Frank Heinemann, Tom Cunningham, Shanker Satyanath, Miriam Golden, Macartan Humphreys, Suresh Naidu, Stephan Meier, Lea Cassar, Gani Aldashev, Maitreesh Ghatak, Ferdinand von Siemens, Elena Paltseva, Jenny Simon, Roland Bénabou and seminar participants at the Berlin Behavioral Economics Workshop, the MPI for Tax Law and Public Finance, Lund University, Queen's University Belfast, Non-Profits, Governments and Organizations Workshop and the Arne Ryde Workshop on Identity, Image and Economic Behavior for their valuable comments and suggestions. Also, I would like to thank Anders Bach-Mortensen for his excellent research assistance. An earlier version of this paper was circulated under the title "Reforming an Institutional Culture."
†WZB Berlin. Contact e-mail: justin.valasek@wzb.eu

1 Introduction

> There are three kinds of [nonmaterial] rewards: a sense of duty and purpose, the status that derives from individual recognition and personal power, and *the associational benefits that come from being part of an organization (or a small group within that organization) that is highly regarded by its members or by society at large.* ∼ James Q. Wilson, *Bureaucracy* (1989) [Emphasis added]

While most economists would agree that state capacity is an important factor for economic performance, there is little consensus on how it can be developed. Arguably, a key element of state capacity is a well-functioning public sector that efficiently allocates resources to public goods and services, and limits distortionary factors such as corruption and elite capture that divert public funds away from socially beneficial outlets. In turn, efficiency within the public sector is optimized when public institutions are staffed by individuals who have a motivation for public service and share the mission (objective) of the institution. This suggests that policies aimed at selecting motivated workers into the public sector offer a promising way to reform underperforming public institutions.

The issue of selection into the public sector and other mission-oriented institutions has received significant attention in recent economic literature: The theoretical literature on public-sector motivation has highlighted the prediction that motivated individuals will disproportionately select into mission-oriented institutions (Francois, 2000 and Besley and Ghatak, 2005, see Francois and Vlassopoulos, 2008 for an overview), and empirical studies in Western Europe have confirmed that public-sector employees exhibit higher relative levels of motivation and altruism than private-sector employees (Gregg et al., 2011 and Dur and Zoutenbier, 2014). However, recent research has provided evidence of reverse sorting in some developing countries with public sectors perceived to be of poor quality: Hanna and Wang (2016) and Banerjee et al. (2015) find that prospective public-sector employees in India are more likely to engage in corrupt behavior than their peers. Moreover, Cowley and Smith (2013) document that, in a subset of developing countries with relatively high levels of corruption, public-sector workers report lower levels of motivation than private-sector workers (see Finan et al., 2015 for an overview of the empirical literature on selection into the public sector).

Given the predictions of the theoretical literature, these new empirical results raise the question of why some public institutions appear to be in equilibria where motivated individuals disproportionately select *out* of employment in the mission-oriented sector. Additionally, given the role selection plays in providing state capacity, an equally important question is how under-performing institutions can be reformed to bring about a transition to an equilibrium with efficient sorting.

In this paper, we show that the puzzle of multiple equilibria can be resolved in a model of labor-market sorting with motivated agents who value both mission and collective reputation, as intimated in the above quote from James Q. Wilson. Additionally, we analyze a dynamic version of the model and highlight the novel prediction that the effect of wages on motivation is conditional on the reputation of the mission-oriented institution. In turn, this implies that a non-monotonic wage path is needed to transition from a low-motivation to a high-motivation steady-state.

The assumption that motivated employees value the collective reputation of their employer is supported by the literature on identity and social image: As argued by Akerlof and Kranton (2005), employees may directly value the identity associated with their job, and will logically seek employment in institutions consistent with their personal identity. Additionally, the collective reputation of an institution can affect worker choice through the channel of prosocial signaling (as in Bénabou and Tirole, 2006, and Ariely et al., 2009) since the collective reputation of an institution signals the type of its employees.[1] Lastly, referring explicitly to bureaucracies, the introductory quote from James Q. Wilson suggests that both mission (purpose) and reputation matter to public-sector employees: while a motivated worker might find a job in a well-regarded non-governmental organization attractive, they might be negatively disposed towards working for a police force widely viewed as corrupt.

To summarize, the model relies on two key assumptions: (i) there exists a motivated type who, all else equal, has a higher productivity in mission-oriented institutions; and (ii) this motivated type values the collective reputation of the institution, and derives positive value from a high reputation and negative value from a low reputation.[2] While we remain agnostic as to the precise mechanism behind this behavioral element, the model we construct is consistent with a micro-foundation based on identity payoffs, prosocial signaling, or homophily.[3]

We first show that the model implies multiple equilibria – both low-motivation equilibria and high-motivation (efficient) equilibria may exist for given parameter values. This multiplicity is intuitive, given that motivated types effectively value an assorta-

[1] A related argument derives from social interaction in the workplace: given a correlation between the collective reputation and workforce composition of an institution, value homophily in the workplace (à la Lazarsfeld and Merton, 1954) provides yet another explanation why workers may value collective reputation. Also, see Henderson and Steen (2015) for an in-depth discussion regarding the role of a firm's reputation in complementing the prosocial-identity of its employees.

[2] Following Tirole (1996), we define collective reputation as the average behavior within the institution. However, in our model, aggregate behavior is a linear function of the proportion of motivated individuals, which allows us to define collective reputation as either aggregate behavior or aggregate work-force composition.

[3] Non-motivated types may also value the collective reputation of the mission-oriented institution because of reputation concerns à la Bénabou and Tirole (2011); our analysis assumes that the motivated type places a greater weight on the collective reputation than the non-motivated type or, equivalently, that both types are characterized by some degree of homophily.

tive labor-market match. Generally, a high-motivation equilibrium is characterized by a higher proportion of motivated types and a *lower* institutional wage relative to the low-motivation equilibrium. The reason a lower wage is maintained in the high-motivation equilibrium is that motivated types are compensated by the high-motivation collective reputation (in addition to any mission payoffs), while the low wage deters non-motivated types from entering the mission-oriented institution.[4] Additionally, this prediction is perfectly consistent with findings of a public-sector wage premium in countries with poor institutional quality.[5]

Several articles have highlighted examples of multiple equilibria in models with motivated agents (see for example Caselli and Morelli, 2004, Macchiavello, 2008, Kosfeld and von Siemens, 2011, and Aldashev et al., 2014). The most novel contribution of our paper is that we formally analyze the problem of reforming a public institution – as far as we are aware, this paper is the first to consider dynamic reform since Jean Tirole's (1996) seminal contribution on collective reputation.[6] Specifically, we ask how a mission-oriented institution with a low reputation and a relatively low proportion of motivated workers can transition to an efficient, high-motivation steady-state equilibrium, using the relative wage as a policy tool. Wages are an effective policy tool since they can be changed transparently and are commonly utilized when instituting public-sector reform (both in theory, e.g. Besley and McLaren, 1993, and in practice, see Rose-Ackerman, 1999 pp. 71-75).

To analyze this question formally, we consider an overlapping-generation version of the model and demonstrate the novel prediction that the effect of a wage change on motivation depends on the initial reputation of the mission oriented institution: starting from a reputation of low-motivation, an increase in the wage *increases* aggregate motivation, while starting from a reputation of high-motivation, an increase in the wage *decreases* motivation. The intuition behind this result lies in the fact that, holding aggregate motivation constant, a wage increase causes an equal proportion of non-motivated and motivated workers to enter the mission-oriented institution. Therefore, if the mission-

[4]Analogous to the efficiency wages in Handy and Katz, 1998, Besley and Ghatak, 2005, and Delfgaauw and Dur, 2007.

[5]Finan et al. (2015) find that public sector wage premiums in many LDCs are positive, and no lower than in developed nations. However, it is unclear whether these premiums are *high enough* to initiate a transition to a high-motivation equilibrium. That is, our model predicts that workers at institutions in a low-motivation equilibrium will be paid a wage premium to offset the reputation cost. Therefore, the findings of the model are perfectly consistent will the low-motivation and relatively high public-sector pay found in many LDCs. As we discuss in the conclusion, a better test of the model come from the comparative statics with respect to the effect of wages on motivation since this prediction of the model is clearly differentiated.

[6]In a recent working paper, Besley and Ghatak (2016) consider the dynamic co-evolution of motivation and rewards structures, and show that societies can converge to different equilibria with different levels of motivation and discretion/control. In contrast, we consider active reform of institutions, and focus on the problem of transition between steady-states.

oriented institution begins with a low average level of motivation, then the proportion of motivated workers increases with a wage increase – leading to an increase in average motivation and making the mission-oriented institution even more attractive to motivated workers. By the same mechanism, however, if the mission-oriented institution begins with a high average level of motivation, than a higher wage decreases average motivation.

This prediction also provides an explanation for the contrasting empirical results regarding the effect of wages on motivation. Specifically, two recent randomized control trials (RCTs) vary wages for mission-oriented institutions and measure the resulting effect on the motivation of applicants: Dal Bó et al. (2013) randomize wages for a position as a community development agent for a program in Mexico and find that higher wages have a *positive* impact on public sector motivation. However, Deserranno (2015) randomizes expected earnings for a position as a community health promoter for a program in Uganda and finds that higher wages have a *negative* impact on prosocial motivation.

Notably, a key difference between the two RCTs is the nature of the employers: while the program in Mexico was directly administered by the Mexican federal government, ranked 103rd out of 175 countries in the Corruption Perceptions Index (Transparency International, 2014), the program in Uganda was administered by BRAC, which is second on the *Global Journal*'s rankings of the top 500 NGOs (2015). Therefore, assuming that the Mexican government has a low reputation while BRAC has a high reputation, the finding that higher wages crowd out motivation in Uganda, but crowd in motivation in Mexico are consistent with the theoretical finding that the effect of wages on motivation depends on the underlying reputation of the mission-oriented institution. This evidence, however, is merely suggestive since the results of the two studies are not directly comparable – as we discuss in the conclusion, a proper test of the model's comparative statics would require controlled variation in reputation and wage within a single population, keeping all other job characteristics constant.

The conditional effect of wages on motivation also implies that, in contrast to the findings of previous papers, a high-motivation steady state equilibrium cannot be achieved by simply setting a low wage. The low wage is a feature, rather than a cause, of the high-motivation equilibrium. Instead, we show that a transitional wage path generally involves an initial increase in the wage to attract more motivated individuals into the mission-oriented institution (crowding in motivation), followed by a gradual decrease of the wage to make the institution less attractive to non-motivated individuals (crowding out non-motivation). The intuition for a non-monotonic transitional wage path follows from the relationship between motivation, wages and reputation outlined above: Starting from a point of low motivation, only an increase in the wage will increase average motivation. Wage increases alone, however, cannot cause a transition to the efficient, low-wage equilibrium – after the level of motivation reaches a threshold level, above which the relationship between motivation and wages switches, the wage must be decreased to effect a transition to the high-motivation, low-wage equilibrium.

We emphasize that the framework we analyze is not peculiar to the public sector and NGOs: to the extent that motivated workers value the collective reputation of generic institutions, the model pertains to any firm or institution that would find it beneficial to attract these workers. For example, firms may seek to replicate the recruiting advantages of, say, Google, whose reputation as a dynamic and attractive employer stems in part from the high quality of its existing workforce; similarly, economics departments may seek to recruit PhD students who are motivated to join academia rather than the private sector, and these academically-motivated students may in turn value a reputation for academic placements.

Crucially, however, we show that a transition from a low reputation to a high reputation is only generally feasible if motivated workers value the mission of the relevant institution. That is, a tipping-point reputation can only be reached through a wage increase if, given a neutral reputation, motivated workers prefer employment in the institution in question over their outside option, as is the case when motivated workers directly value the social output of a public institution (i.e. mission-contingent payoffs à la Besley and Ghatak, 2005). This finding suggests that transitions are not feasible in generic institutions and that achieving a transition may require a firm to actively invest in, say, corporate social responsibility,[7] or that transitions are only possible for departments at universities with an overall reputation for academic excellence.

Lastly, we consider the case of correlation between a worker's ability and their level of motivation, and show how this correlation can be leveraged to achieve a transition. This case also provides insights into a commonly-attempted strategy of creating "elite" divisions within an existing institution – for this strategy to be successful, the institution must both recruit disproportionately from an ability type with a high average level of motivation *and* offer a relatively high wage. Without a high initial wage, the strategy of recruiting from a high-motivation ability type may not be sufficient, since the overall low reputation of the institution creates an adverse selection problem.

1.1 Literature

In the classic study cited at the beginning of this paper, Wilson remarks that, given the lack of incentives... "what is surprising is that bureaucrats work at all" (1989). More generally, it has been argued that non-monetary incentives in the workplace play an important role in determining worker's behavior (Dewatripont et al., 1999, Akerlof and Kranton,

[7]The management literature suggests that corporations engage in charitable activities for precisely this purpose; see for example Bhattacharya et al. (2008) "Using Corporate Social Responsibility to Win the War for Talent."

2000, Akerlof and Kranton, 2005, Prendergast, 2008, Huck et al., 2012, and Fischer and Huddart, 2008). A subset of this literature considers motivation in the workplace, and has largely focused on optimal contracting in the presence of a motivated type, given that non-monetary incentives can be crowded out or distorted by traditional monetary incentive contracts (e.g. Murdock, 2002, Dixit, 2002, Sliwka (2007) and Ellingsen and Johannesson, 2008; see Francois and Vlassopoulos, 2008 and Prendergast, 2008 for an overview).

Another strand of this literature, in which our paper arguably falls, is concerned with the question of optimal contracting with endogenous worker sorting into the mission-oriented sector, given the presence of different behavioral types. Several papers highlight that the efficiency wage in the mission-oriented sector should be low relative to the private sector, as a low wage will disproportionately attract workers who have public sector motivation and who are compensated by non-pecuniary benefits of public-sector employment (Handy and Katz, 1998, Francois, 2000, and Besley and Ghatak, 2005). This result has been extended to account for the fact that other facets of the public sector may disproportionately attract individuals with harmful qualities, such as laziness or antisocial motives (Delfgaauw and Dur, 2008, Auriol and Brilon, 2014), or into positions where altruism is counter-productive (Prendergast, 2007). In an article contemporary to this paper, Henderson and Steen (2015) consider a model of motivated agents who signal their prosocial identity through the reputation of their employer, where reputation is a function of the actions of the firm, and show how this can cause firms to benefit from endogenously choosing a prosocial purpose (mission).

Our main contribution to this literature is that by considering motivated agents who value the collective reputation of an institution, the question of optimal contracting is transformed from a problem of static equilibrium selection to a problem of dynamic transition, since collective reputation functions as a state variable. That is, similar to the situation analyzed in Tirole (1996), the institution and its current workers are burdened with the legacy of past behavior, which implies that the impact of incentives becomes sensitive to the institution's starting point: higher wages increase motivation in a low-motivation institution, but decrease motivation in a high-motivation institution. Therefore, reforming, say, a culture of corruption requires a more complex approach than simply replicating the incentives of a low-corruption institution.

Our paper is also related to a set of papers that analyze transitions between norms (e.g. Bisin and Verdier, 2001, Besley et al., 2014, Bidner and Francois, 2013, and Acemoglu and Jackson, 2015). For example, in a paper related to ours in spirit, Acemoglu and Jackson (2016) detail how endogenously enforced laws can be changed dynamically to transition between a steady-state of lawlessness to a steady-state of law-abiding. They show that a sudden shift in laws away from the current norm of behavior can be counter-productive, but that a series of incremental shifts can result in a transition. Our paper also shows that the path of a reform matters – a transition cannot be enacted by skipping straight to the wage of the efficient equilibrium. However, in contrast to the findings of Acemoglu and Jackson (2016), we find that the policy tool we consider (wages) must take a non-monotonic path for the system to transition to the optimal steady state.

Lastly, we argue that our results help reconcile the well-known policy prescription of a high public-sector "efficiency wage" to deter corruption (Besley and McLaren (1993)) with the concern that higher wages will crowd out intrinsic motivation (Besley and Ghatak, 2005). Indeed both results have empirical support: higher public-sector wages are weakly correlated with lower corruption (Treisman, 2000, Van Rijckeghem and Weder, 2001, and Di Tella and Schargrodsky, 2003), and there is evidence for a below-market "public-sector efficiency wage" (see Gregg et al., 2011) in developed nations. The theoretical results of this paper help reconcile these empirical findings: In contexts where public institutions have a good reputation, higher wages will simply crowd out motivated workers. However, in contexts where public institutions have a poor reputation, for example, due to high levels of corruption, then wage increases help combat corruption directly through the efficiency-wage argument, and indirectly by increasing the average motivation of public-sector workers.

2 Static Model

In this section we introduce a simple model that illustrates the relevant results. We set up the model from the perspective of a mission-oriented institution, A.

Agents:

There is a continuum of workers of measure one with a compact index set I. Workers are one of two types: Non-motivated or Motivated. Take $m_i = 1$ if worker i is motivated and $m_i = 0$ if non-motivated. A proportion λ of workers are motivated. Workers have an ability y_i, and an outside option of x_i, which can be thought of as worker i's wage in their outside-option employment. For simplicity, we constrain $y_i = 1$ for the main analysis (this assumption is relaxed in Section 5.2), while x_i is heterogenous and distributed according to a commonly-known uniform distribution with support $[\underline{x}, \overline{x}]$.[8] That is, all agents have the same inherent ability at the mission-oriented institution, but vary in their outside option employment opportunity. Additionally, x_i is uncorrelated with worker motivation, although we also relax this assumption in Section 5.2.

Take $a_i = 1$ if worker i is employed in institution A, and $a_i = 0$ if i is employed in

[8]The assumption of a uniform distribution of ability significantly simplifies the analysis, and allows us to characterize the comparative statics of the dynamic model. However, the main results of the analysis, namely multiple equilibria and the existence of a dynamic transition path, do not depend on this assumption (we discuss this in more detail in footnote 17).

their outside option.

Payoffs:
Institution A has a demand for labor of measure $\nu \leq \min\{\lambda, (1-\lambda)\}$, and receives the following the profit, or net social output, from each individual it hires:

$$\pi_i^A = \pi y_i + \beta \mathbb{1}(m_i = 1) - w_i$$

Where y_i is the ability of worker i, β reflects the higher productivity of motivated workers at institution A (we discuss the foundation for this modeling feature below), and w_i is the wage paid to worker i. Institution A does not observe π_i^A directly, but aggregate profit, $\pi^A = \int_I \pi_i^A a_i$, is publicly observable. Importantly, each worker's outside option, x_i, is private information and is unobserved by A.[9] Therefore, from A's perspective, all workers are ex ante identical, and we assume institution A offers a constant wage, w, to all workers – while we impose a uniform wage here for simplicity, in the Appendix we show that a uniform wage is always optimal (see Result 1 in Section 6.1).

We add one assumption regarding institution A's profit function, relative to ν. Take x' that solves:

$$(1-\lambda)\frac{x' - \underline{x}}{\overline{x} - \underline{x}} = \nu. \tag{1}$$

Assumption 1
The output that institution A receives from each non-motivated worker, π, is greater than $\nu + x'$.

This assumption ensures that, holding average worker motivation constant, institution A always maximizes profit at a wage that ensures full employment.

Definition 1 (Collective Reputation)
The collective reputation of institution A is equal to $C = \int_I a_i m_i / \int_I a_i$.

We define the collective reputation using the proportion of motivated types in institution A instead of using the aggregate behavior within A; however, the two are equivalent here since type is perfectly correlated with behavior. Therefore, this definition is consistent with reputation payoffs that depend on type-composition (social signaling and homophily) or institutional behavior (identity). In other words, agents can infer the composition of types within institution A by observing A's aggregate social output. Since the collective reputation of the mass of individuals who receive their outside option is perfectly nega-

[9]The assumption that A does not perfectly observe x_i is a key assumption of the model, since it rules out the possibility of A offering an individualized wage contract that conditions on x_i, which would allow A to perfectly screen workers for motivation by offering worker i a wage that is slightly less than i's outside option (see Besley and Ghatak, 2005). In Section 4.4 we analyze an extension of the model where workers have a publicly-observable type that is correlated with both motivation and the outside option.

tively correlated with C, we can interpret C as A's relative reputation, and we do not explicitly consider reputation payoffs associated with the outside option.

Non-motivated workers have a standard linear utility function over income:

$$u_n(w_i) = w_i.$$

Where w_i is equal to w if i is employed in A, and equal to x_i if i receives the outside option.

Motivated workers differ from non-motivated workers in three regards: (1) they are more productive if matched with institution A, (2) they may value the mission of institution A, and hence may receive a direct benefit from employment in institution A (as in Francois, 2000 and Besley and Ghatak, 2005), (3) they value the collective reputation (workforce composition) of the mission-oriented sector A, e.g. due to type signaling or a direct preference for workplace homogeneity. To reflect both (2) and (3), motivated workers have a utility function of the following form:

$$u_m(w_i, C, a_i) = w_i + v(C)a_i.$$

Where C is the proportion of motivated workers in institution A, and $v(C)$ captures motivated workers' payoffs from both collective reputation and mission; $v(\cdot)$ is strictly increasing and concave. In the analysis, we also utilize the following general utility function:

$$u(w_i, C, a_i, m_i) = w_i + v(C)a_i m_i.$$

To be consistent with the intuition that motivated workers receive a positive payoff from a high reputation, and a negative payoff from a low reputation, we restrict the analysis to $v(1) > 0$ and $v(0) < 0$. Additionally, the main analysis focuses on the case of $v(\lambda) > 0$. Given a "generic" firm, it may be natural to assume that, given a neutral reputation, a motivated agent does not receive a positive relative payment from employment in A; i.e. that $v(\lambda) = 0$. However, in this paper we primarily focus on "mission-oriented" institutions, where motivated agents are directly motivated by the mission, or product, of institution A (as in Besley and Ghatak 2005). In our model, given the constant product produced by each worker in institution A, mission-motivation simply translates into a constant benefit of working for institution A. Therefore, holding constant reputation and wage between sectors, motivated workers prefer working at the mission-oriented sector, which implies that $v(\lambda) > 0$, where $v(\lambda)$ represents the mission-benefits a motivated worker receives from employment in the mission-oriented institution.[10]

Since we are considering the wage of institution A as a policy tool, it is necessary

[10] In Section 4.2, we characterize the case of a generic firm ($v(\lambda) = 0$) and show that mission-orientation is a necessary condition for reform.

to specify the framework for employment in institution A when it is over-demanded (i.e. demand for employment is greater than ν). Since all workers are ex-ante identical from A's perspective, workers are randomly selected for employment in institution A from among the applicants.

Formally, workers choose $\hat{a}_i \in \{0,1\}$ at the beginning of the period, which determines employment according to the following rule:

$$a_i \begin{cases} = 0 & \text{if } \hat{a}_i = 0 \\ = 1 \text{ w.p. } q & \text{if } \hat{a}_i = 1. \end{cases}$$

Where:

$$q = \min\left\{1, \frac{\nu}{\int_I \hat{a}_i}\right\}.$$

OPTIMALITY AND EQUILIBRIUM:
Throughout the paper, we consider the objective of maximizing the efficiency of the mission-oriented sector, represented by social output, π^A.[11] Therefore, similar to the optimal-contracting approach of Besley and Ghatak (2005), we define outcomes given a wage in the mission-oriented institution, rather than analyzing w as an equilibrium choice.

Since information is complete, the equilibrium concept we use is Nash. That is, an equilibrium is defined by a set of employment choices, $\{\hat{a}_i\}$, such that given w and C, non-motivated workers set $\hat{a}_i = 1$ iff $u_n(w, C, a_i = 1) = w \geq x_i$, non-motivated workers set $\hat{a}_i = 1$ iff $u_m(w, C, a_i = 1) = w_i + v(C) > x_i$, and $C = \int_I \hat{a}_i m_i / \int_I \hat{a}_i$.[12]

3 Analysis of the Static Model

In the analysis, we will use the terminology High/Low-Motivation to classify equilibria:

Definition 2 (Collective Reputation High/Low-Motivation)
The collective reputation of institution A is High-Motivation *if $C > \lambda$ and* Low-Motivation *if $C \leq \lambda$.*

That is, the institution has a high-motivation reputation if the proportion of motivated workers in institution A is greater than the population average. Additionally, we refer to an equilibrium with a high-motivation (low-motivation) reputation as a high-motivation (low-motivation) equilibrium.

[11] However, note that given full employment in the mission-oriented sector, equilibria with greater π^A (higher C) also produce greater aggregate utility, since both output in the mission-oriented institution and workers' reputation-payoffs are higher.

[12] A more precise definition of equilibrium would include the set of employment outcomes, $\{a_i\}$; however, to allow for more efficient notation, we use the fact that with a continuum of workers, C can be defined as a function of $\hat{a}_i$.

First, we characterize equilibria in the static model in terms of the cutoff types, x_m for motivated workers and x_n for non-motivated workers:

Lemma 1 (Cutoff Equilibrium)
Given w, equilibrium employment decisions are characterized by $\{x_m, x_n\}$, where $\hat{a}_i = 1$ if and only if $m_i = 1$ and $x_i \leq x_m$ or $m_i = 0$ and $x_i \leq x_n$.

Lemma 1 states that, in equilibrium, conditional on type, individuals with a relatively low outside option will select into institution A. The result follows from single-crossing, by motivation-type, of $u(x_i, C, a_i = 1, m_i)$ and $u(x_i, C, a_i = 0, m_i)$ in x_i. (Formal proofs of all results can be found the Appendix.)

Lemma 1 also allows us to characterize equilibria by identifying the individuals who are indifferent between the mission-oriented institution and their outside option, as well as the corresponding C. That is, interior equilibria are defined by $\{x_m, x_n, C\}$ that solve the following system of equations:

$$\begin{aligned} x_m &= w + v(C), \\ x_n &= w, \\ C &= \frac{\lambda(x_m - \underline{x})}{(1-\lambda)(x_n - \underline{x}) + \lambda(x_m - \underline{x})}. \end{aligned}$$

To simplify this notation, we define the function $C(x, w)$ as:

$$C(x, w) = \frac{\lambda(x - \underline{x})}{(1-\lambda)(w - \underline{x}) + \lambda(x - \underline{x})},$$

which allows us to define interior equilibria as $\{x_m, x_n, C(x_m, w)\}$ such that $x_m = w + v(C(x_m, w))$. Moreover, this allows us to define $u_m(x_i, C(x_i, w), a_i = 1)$, which is the utility a motivated type, i, receives from employment in A *if* all motivated workers with $x_j \leq x_i$ and all non-motivated workers with $x_j \leq w$ are employed in A, and all others take their outside option. Since we define equilibria given w, when convenient we refer to $C(x, w)$ as a function of x_m only ($C(x_m)$) and further simplify the notation by referring to an interior equilibrium as an intersection of $u_m(w, C(x_m), a_i = 1) \equiv u_m(w, C(x_m))$ and x_m.

Given these definitions, we show that a high-motivation equilibrium exists for all w.

Lemma 2 (Existence high-motivation equilibrium)
Given $w \in (\underline{x}, \overline{x})$, there exists a unique high-motivation equilibrium, $\{x_m, x_n, C^h\}$, with $C^h > \lambda$.

Existence follows from mission-motivation: since $v(\lambda) > 0$, this implies that, given $x_n = w$, either an interior crossing of $u_m(w, C(x_m))$ and x_m exists, or the corner $x_m = \overline{x}$ is

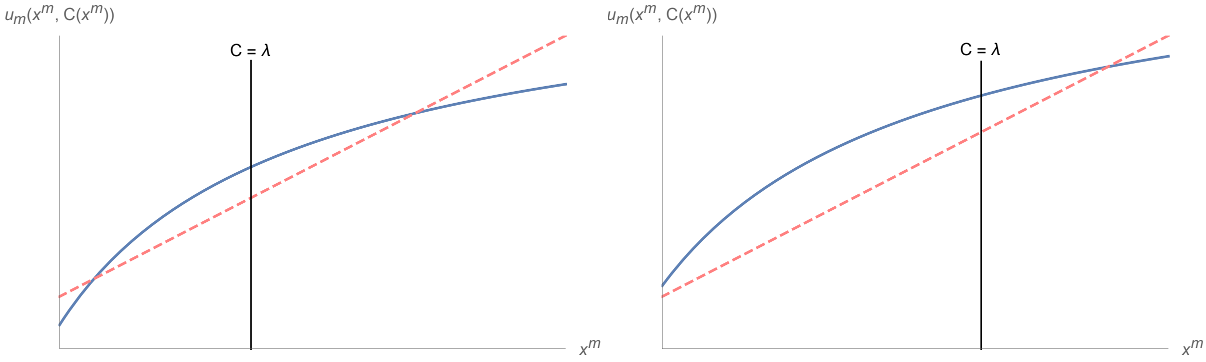

Figure 1: These graphs illustrate the respective utility of employment in A (blue, solid line) and the outside option (red, dashed line) for a motivated type with $x_i = x_m$, *given* that all motivated workers with $x_i < x_m$ set $\hat{a}_i = 1$. The left graph has $w + v(0) < \underline{x}$, and the right graph has $w + v(0) > \underline{x}$.

an equilibrium. Since both $v(\cdot)$ and $C(x_m)$ are concave, $u_m(w, C(x_m))$ is concave in x_m, which implies uniqueness.

Figure 1 illustrates equilibria for a given value of w; the graph illustrates the respective utility of employment in A and the outside option for a motivated type with $x_i = x_m$, *given* that all motivated workers with $x_i < x_m$ set $\hat{a}_i = 1$. Again, interior equilibria are represented by intersections of $u_m(w, C(x_m))$ and x_m since at an intersection, motivated workers with $x_i = x_m$ are indifferent between employment in institution A and their outside option.

Proposition 1 (Condition for Unique Equilibrium/Multiple Equilibria)
If $w > \underline{x} - v(0)$ or $w \in (\underline{x} - v(1), \underline{x})$ then the high-motivation equilibrium is the unique equilibrium of the model. If $w \in (\underline{x}, \underline{x} - v(0))$, then both high-motivation and low-motivation equilibria exist.

A graphical proof of Proposition 1 follows from Figure 1: If $w > \underline{x} - v(0)$, then $u_m(w, C(x_m))$ is greater than x_m at $\underline{x}$ and λ. Therefore, by the concavity of $u_m(w, C(x_m))$ in x_m, there is no crossing of $u_m(w, C(x_m))$ and x_m in the range $[\underline{x}, \lambda]$ (right graph). If $w < \underline{x} - v(0)$, however, then $u_m(w, C(x_m))$ is smaller than x_m at $\underline{x}$ and larger than x_m at λ. Therefore, a crossing of $u_m(w, C(x_m))$ and x_m must exist in the range $(\underline{x}, \lambda)$ (left graph).

While Proposition 1 characterizes the existence of high/low equilibria as a function of the wage in institution A, it is of limited relevance when considering the optimal equilibrium from the perspective of institution A. Instead, market-clearing equilibria, where $\int_I \hat{a}_i = \nu$, will be a relevant benchmark since, by Assumption 1, given a fixed level of average motivation, a market-clearing equilibrium is optimal from the perspective of the mission-oriented institution. Therefore, the following result will be useful when characterizing the optimal equilibrium.

Lemma 3 (Market-Clearing Equilibria)
(i) A high-motivation, market-clearing equilibrium $\{w^h, C^{h'}\}$ *exists.*
(ii) If $\nu < (1-\lambda)|v(0)|$*, then a low-motivation market-clearing equilibrium* $\{w^l, C^{l'}\}$ *exists with* $w^l > w^h$*.*

The existence of a high-motivation, market-clearing equilibrium follows from the fact that the cutoff abilities in the unique high-motivation equilibrium, $\{x_m, x_n\}^h$, are both continuous functions of w, which implies that an equilibrium with $\int_I \hat{a}_i = (1-\lambda)(x_n-\underline{x})+\lambda(x_m-\underline{x}) = \nu$ exists. While low-motivation market-clearing equilibria do not exist for all parameter values, Lemma 3 provides a sufficient condition and shows that they are present when $|v(0)|$ is sufficiently high. Additionally, note that a low-motivation market-clearing equilibrium requires a higher wage than the high-motivation market-clearing equilibrium since, for motivated workers, additional monetary compensation is needed to offset the low reputation payoffs.

3.1 Optimal Contract

Next, we will partially characterize the optimal equilibrium from the point of view of the mission-oriented institution. Note that even though a market-clearing wage is optimal given a fixed level of average motivation (by Assumption 1), as the following proposition illustrates, the equilibrium that maximizes $\pi^A(x_m, x_n, w)$ need not coincide with a market-clearing equilibrium.

Proposition 2 (Optimal Equilibrium)
The equilibrium that maximizes $\pi^A(x_m, x_n, w)$*,* $\{w^*, C^*\}$*, is either equal to a market-clearing, high-motivation equilibrium,* $\{w^h, C^{h'}\}$*, or satisfies* $w^* \leq w^h$*,* $C^* \geq C^{h'}$*.*[13]

First, note that Proposition 2 implies that the optimal equilibrium must be a high-motivation equilibrium. This follows from the fact that for any interior low-motivation equilibrium, there exists an equivalent high-motivation equilibrium with higher net social output. Second, Proposition 2 states that $\{w^*, C^*\}$ need not correspond to a market-clearing equilibrium: since full employment is optimal for A given a fixed level of reputation, $w^* < w^h$ only if $\int_I \hat{a}_i^* m_i > \int_I \hat{a}_i^h m_i$. Formally, it is possible for $w^* < w^h$ since x_m can be a decreasing function of w – therefore, an equilibrium with a below-market wage can be optimal for A, since it may feature a greater *absolute* number of motivated workers.

While Proposition 2 identifies the optimal equilibrium, the existence of multiple equilibria implies that the model may be indeterminate regarding the optimal contract. Specifically, for $w^* > \underline{x}$, Proposition 1 demonstrates that both high and low-motivation

[13]Note that this proposition characterizes the optimal equilibrium with a uniform wage; however, in the Appendix we show that $\{w^*, C^*\}$ remains optimal even when A is able to offer different wages to different subsets of workers (see Result 1).

equilibria exist for w^* when $|v(0)|$ is large, which implies that the mission-oriented institution cannot ensure optimality by simply setting wages equal to w^* (for a detailed discussion of the indeterminacy of policy analysis in the presence of multiple equilibria, see Morris and Shin, 2000).

It is also important to note that while Proposition 1 shows that the mission-oriented institution can always solve the problem of equilibrium multiplicity by setting a sufficiently high wage, such a solution is costly, since it requires a wage higher than w^*. In fact, depending on parameter values, setting a wage high enough to ensure that the high-motivation equilibrium is unique may leave the mission-oriented institution *worse off* then in the low-motivation market clearing equilibrium. The static model also suggests that the mission-oriented institution can achieve a high motivation equilibrium by setting a wage below $\underline{x}$. However, this solution is also costly for $w^* > \underline{x}$, since it requires a wage at which employment in A is smaller than ν. Additionally, as we will show in the following section, in a dynamic setting where reputation is persistent, decreasing wages to below $\underline{x}$ will not result in a transition to a high-motivation equilibrium if A is endowed with an initial low-motivation reputation.

Therefore, instead of relying on criteria for equilibrium selection, we instead consider the problem of transitioning from a point of low-motivation to the optimal high-motivation equilibrium. Arguably, characterizing such a transition is a first-order concern, since the persistence of collective reputation (see Tirole, 1996), combined with legal and transactional constraints that may prevent institutions from immediately replacing their entire workforce, makes it unlikely that institutions in a low-motivation equilibrium will suddenly shift to a high-motivation equilibrium. Therefore, in the next section, we take a similar approach to Tirole (1996) and Acemoglu and Jackson (2016) and introduce a dynamic version of the model that accounts for the friction stemming from the persistence of reputation, and that allows us to characterize a path for dynamically reforming a public institution.

4 Dynamic Model

We now add a dynamic layer to the static framework, and consider an overlapping-generation (OLG) framework with an infinite time horizon. We also introduce a plausible source of friction to the dynamic model, namely that motivated workers value the *lagged* collective reputation of institution A. This captures the notion that reputations and reputation payoffs are sticky, as perceptions might not update automatically (as in Besley et al., 2014, and Tirole, 1996, who argues that "...stereotypes are long-lasting because new members of a group at least partially inherit the collective reputation of their elders"). Most importantly, this friction implies that the transitions from low to high-reputation illustrated in this paper do not rely on the coordinated action of a mass of motivated workers.[14]

As before, in each period there is a continuum of workers of measure one. Following the standard OLG approach (analogous to Acemoglu and Jackson, 2015), however, workers live for two periods, with half of the workers "born" in the current period and half born in the period prior. We denote workers by (i,k), where $k \in \{1,2\}$ denotes whether i is in his first or second period of life (each generation has a compact index set I). The distribution of $m_{(i,k)}$ and $x_{(i,k)}$ for each generation is identical to that of the static framework (a proportion λ are motivated and $x_{(i,k)} \sim [\underline{x}, \overline{x}]$).

Formally, workers have period-utility payoffs analogous to the static model, with the exception that $v(\cdot)$ is a function of C^{t-1}:

$$u^t(w^t_{(i,k)}, C^{t-1}, a^t_{(i,k)}, m_{(i,k)}) = w^t_{(i,k)} + v(C^{t-1})a^t_{(i,k)}m_{(i,k)}$$

Where $C^t = \big(\sum^k \int_I a^t_{(i,k)}m_{(i,k)}\big)/\big(\sum^k \int_I a^t_{(i,k)}\big)$ is equal the proportion of motivated workers in institution A in period t. Given that motivated workers value last period's reputation, workers' period payoffs are not a function of their expectations regarding the proportion of motivated workers that will enter institution A's workforce in the current period. However, since expectations regarding future periods enter dynamic payoffs, the equilibrium path of $\{C^t\}$ is not independent of expectations.

The workers' employment contracts last for both periods of the worker's life. That is, in the first period of their lives workers observe the wage offered in institution A, w^t, and choose $\hat{a}^t_{(i,1)} \in \{0,1\}$. Young workers matched with their outside option employment are also matched with their outside option employment in $t+1$. Young workers matched with institution A, however, retain their outside option in their second period of life; that is, in period $t+1$ these workers choose between remaining employed in A with wages w^t, or exiting A and receiving their outside option. This structure, which assumes binding tenure and fixed wages, simplifies the analysis and is consistent with the relatively rigid nature of wages and tenure observed in many public bureaucracies (see the discussion in Finan et al., 2015 pp. 2-3). Moreover, a previous version of this paper (Valasek, 2016) demonstrates that the findings of the current paper are robust to the assumption that wages and employment is "fixed" for the older generation of workers.[15]

To keep the notation consistent, we represent the choice of employment in the second period by $\hat{a}^t_{(i,2)} \in \{0,1\}$; however, $\hat{a}^t_{(i,2)}$ is only a choice variable for workers with $a^t_{(i,1)} = 1$.

[14] Additionally, $C^t = C^{t-1}$ if employment in A is zero in time t – this assumption rules out transition paths where institution A 'resets' its collective reputation by choosing a wage low enough such that employment is equal to zero.

[15] Instead of an overlapping generations model, Valasek (2016) utilizes a dynamic model with a "Poisson death process," similar to Tirole (1996), where agents choose enter or exit institution A in each period, and where wages are allowed change even for workers with tenure.

Analogous to the static model, workers' first-period employment is determined according to the following rule:

$$a^t_{(i,1)} \begin{cases} = 0 & \text{if } \hat{a}^t_{(i,1)} = 0 \\ = 1 \text{ w.p. } q & \text{if } \hat{a}^t_{(i,1)} = 1 \end{cases}$$

Where:

$$q = \min\left\{1, \frac{\nu}{\int_I a^t_{(i,2)} + \int_I \hat{a}^t_{(i,1)}}\right\}.$$

That is, if the mission-oriented institution is over-demanded, "open" slots in A are randomly allocated to new applicants. As detailed above, workers' second-period employment is determined according to the following rule:

$$a^t_{(i,2)} \begin{cases} = 0 & \text{if } a^t_{(i,1)} = 0 \text{ or } \hat{a}^t_{(i,2)} = 0 \\ = 1 & \text{if } a^t_{(i,1)} = 1 \text{ and } \hat{a}^t_{(i,2)} = 1. \end{cases}$$

The timing of the period game is as follows:

1. A sets w^t and w^t, C^{t-1} is observed by workers.
2. Young workers choose $\hat{a}^t_{(i,1)} \in \{0,1\}$ and, conditional on $a^{t-1}_{(i,2)} = 1$, old workers choose $\hat{a}^t_{(i,2)} \in \{0,1\}$.
3. Period utility $(a^t_{(i,k)})$ realizes.

Dynamic payoffs:

Institution A has period payoffs that are equivalent to the static model, and dynamic payoffs equal to:

$$\Pi_A = \sum_t^{\infty} (1-\tau)^{t-1} \pi^t_A.$$

Where $\tau \in [0,1]$.

For workers, the dynamic setting introduces the possibility of a positive option value of employment in institution A. Therefore, a young workers' utility of employment in A takes the following form:

$$U(w^t, \{C^t\}, m^t_{(i,1)}, x^t_{(i,1)}, a^t_{(i,1)} = 1, a^{t+1}_{(i,1)}) = u^t(w^t, C^{t-1}, a^t_{(i,1)} = 1, m^t_{(i,1)}) + \delta u^{t+1}(w^t, C^t, m^t_{(i,1)}, x^t_{(i,1)}),$$

where $\delta \in [0,1]$. Note that since (i,k) retains his outside option at $k=2$, we define u^{t+1} as follows:

$$u^{t+1}(w^t, C^t, m^t_{(i,1)}, x^t_{(i,1)}) = max\{u^t(w^t, C^t, m^t_{(i,1)}, a^{t+1}_{(i,1)} = 1), x_{(i,1)}\}.$$

Lastly, for agents matched with their outside option in their first period of life, $U(w^t, \{C^t\}, m^t_{(i,1)}, x^t_{(i,1)}, a^t_{(i,1)} = 0, a^{t+1}_{(i,1)}) = (1+\delta)x^t_{(i,1)}$, and for $k=2$ the dynamic utility is simply equal to the period utility.

Since we are concerned with reforming an existing institution, we consider the situation where A "inherits" a reputation and workforce. For the purpose of illustration, we consider the case where institution A is endowed with a reputation, C^0, and wage, w^0, that correspond to a market-clearing low-motivation static equilibrium (below we show that static equilibria correspond to steady-state equilibria).[16] This assumption is made for illustrative purposes only and does not impact the transitional wage path illustrated in the following section.

EQUILIBRIUM:
Since information is complete, we use sub-game perfect Nash Equilibrium. Additionally, we follow the selection-criterium of Gul et al. (1986) and assume that agents do not condition their choices on the past actions of sets of agents of measure zero, which insures that unilateral deviations by a single worker do not affect the actions of other workers. Given a set of wages $\{w^t\}_0^\infty$ and an initial reputation C^0, an equilibrium constitutes a set of employment choices, $\{\hat{a}^t_{(i,k)}\}_1^\infty$, that maximize each worker's dynamic utility given the implied reputation, $\{C^t\}_0^\infty$.

Similar to the static section, we consider the objective of transitioning institution A to the steady-state equilibrium that maximizes net social output (we define steady-state equilibria formally below). While this objective is not always equivalent to maximizing the discounted stream of profits, as long as institution A has a discount rate, τ, that is low enough, then a wage path that transitions to the steady-state that maximizes social output will always be preferable to remaining in any other steady state. We address cost-minimizing transitions after introducing our main result.

4.1 Analysis of Dynamic Model

We introduce some general results before addressing the issue of transition. First, we characterize an individual's decision rule, fixing $\{w^t\}$ and $\{C^t\}$. Each worker chooses $\hat{a}^t_{(i,1)} = 1$ if, and only if, the following expression holds:

$$u^t(w^t, C^{t-1}, m_{(i,1)}) + \delta u^{t+1}(w^t, C^t, m_{(i,1)}, x_{(i,1)}) \geq (1+\delta)x_{(i,1)}. \tag{2}$$

Note that since $(i,1)$ can always choose their outside option in their second period of life, they will choose $\hat{a}^t_{(i,1)} = 1$ if $u^t(w^t, C^{t-1}, m_{(i,1)}) \geq x_{(i,1)}$. The inverse statement, however, is not always true: a motivated worker $(i,1)$ may choose $\hat{a}^t_{(i,1)} = 1$ when $u^t(w^t, C^{t-1}, m_{(i,1)}) < x_{(i,1)}$ given the expectation that the reputation of institution A will increase in the current period, making employment in A preferable to the outside option in $t+1$. Therefore,

[16] Additionally, A has $t = 0$ workforce such that $\int_I a^0_{(i,1)} = \nu/2$, and $a^0_{(i,1)} = 1$ iff $m_{(i,1)} = 1, x_{(i,1)} < x^0_m$ and $a^0_{(i,1)} = 1$ iff $m_{(i,1)} = 0, x_{(i,1)} < x^0_n$, where the cutoff types $\{x^0_m, x^0_n\}$ correspond to the equilibrium cutoffs of the low-motivation static equilibrium.

a motivated worker's decision rule at $k = 1$ is dependent on the decision rule of other motivated workers, which can result in multiple equilibria for any set of wages $\{w^t\}$ (equilibria are not "expectations free").

The following lemma extends the result of Lemma 1 to the dynamic model:

Lemma 4 (Cutoff Equilibrium)
Given $\{w^t\}$, equilibrium employment decisions are characterized by a set of cutoff values $\{x^t_{m,k}, x^t_{n,k}\}$, where $\hat{a}_{(i,k)} = 1$ if and only if $m_{(i,k)} = 1$ and $x_{(i,k)} \leq x^t_{m,k}$, or $m_{(i,k)} = 0$ and $x_{(i,k)} \leq x^t_{n,k}$.

First, note that in period t the arguments of the utility function of the second generation, specifically C^{t-1} and w^{t-1}, are fixed, which implies that the equilibrium cutoff values $x^t_{m,2}$ and $x^t_{n,2}$ are unique. Accordingly, the bulk of our analysis will focus on decision of young workers in period t. Therefore, to simplify the notation in the following analysis, we will drop the k subscript on all variables when referring to variables pertaining to generation $k = 1$: e.g. $x^t_m \equiv x^t_{m,1}$ and $\hat{a}^t_i \equiv \hat{a}^t_{(i,1)}$.

Of course, the employment decisions of the older workers still impact the employment decision of young workers in period t through C^t. However, in order to present the results of the dynamic model as simply as possible, instead of directly referring to the employment decisions of the older generation, we introduce the following decomposition of C^t that implicitly accounts for the employment decisions of $k = 2$:

$$C^t = \phi^t_1 C^t_1 + \phi^t_2 C^t_2,$$

where C^t_k is equal to the proportion of generation k workers in A that are motivated, $\int_I a_{(i,k)} m_{(i,k)} / \int_I a_{(i,k)}$, and ϕ^t_k is equal to the proportion of workers in A that are of generation k, $\int_I a_{(i,k)} / \sum_k \int_I a_{(i,k)}$. Since C^t implicitly includes the employment decision of the older generation, we simplify our notation of an equilibrium to $\{x^t_m, x^t_n, C^t\}^\infty_t$.

The main focus of our analysis is on the optimal steady-state equilibrium, where steady-state equilibria are defined as follows:

Definition 3 (Steady-State Equilibria)
Given $\{w^t\}$ such that $w^t = \bar{w}$ for all $t \geq t'$, an equilibrium $\{\bar{w}, x^t_m, x^t_n, C^t\}^\infty_{t'}$ is a steady-state equilibrium if $x^t_m = \bar{x}_m$, $x^t_n = \bar{x}_n$, and $C^t = \bar{C}$ for all $t \geq t'$.

The following proposition clarifies the relationship between static and dynamic equilibria:

Proposition 3 (Static Equilibrium $\Leftrightarrow$ Steady-State Equilibrium)
For each static equilibrium, $\{w, x_m, x_n, C\}$, there exists a corresponding steady-state equilibrium, $\{\bar{w}, \bar{x}_m, \bar{x}_n, \bar{C}\}^\infty_{t'}$, with $\bar{w} = w$, $\bar{x}_m = x_m$, $\bar{x}_n = x_n$, and $\bar{C} = C$; and for each steady-state equilibrium, there exists a corresponding static equilibrium.

Proposition 3 shows that when both high and low-reputation equilibria exist in the static model, then corresponding high and low-reputation steady-state equilibria exist in the dynamic model. Additionally, it gives the following corollary:

Corollary 1
The optimal stead-state equilibrium $\{\bar{w}^*, \bar{C}^*\}_t^\infty$*, corresponds to the optimal static equilibrium,* $\{w^*, C^*\}$*.*

In the following text, we use $\{w^*, C^*\}$ to refer to the optimal steady-state equilibrium.

Next we introduce two lemmas that will allow us to drastically simplify the analysis of a dynamic transition.

Lemma 5
Given a wage path $\{w^t\}$ *and* $\delta = 0$*, the dynamic equilibrium is unique.*

Note that $\delta = 0$ implies that workers in both their first and second periods of life will make employment decisions myopically, only setting $\hat{a}^t_{(i,k)} = 1$ if $u^t(w^t, C^{t-1}, m_{(i,k)}) > x_{(i,k)}$. Therefore, the employment decision is a function of variables that are fixed in period t, which implies uniqueness given the initial workforce and reputation at $t = 0$. Lemma 5 allows us to define the myopic equilibrium as follows:

Definition 4 (Myopic Equilibrium)
Given $\{w^t\}$*, take the myopic equilibrium,* $\{w^t, \tilde{x}^t_m, \tilde{x}^t_n, \tilde{C}^t\}$*, to be equal to the unique equilibrium that follows when* $\delta = 0$*.*

The non-uniqueness of equilibria in the general case with $\delta > 0$ makes it difficult to directly analyze the impact of wage reform. However, the following lemma links the myopic equilibrium to the dynamic equilibria of the game with $\delta > 0$, and will allow us to partially characterize *all* dynamic equilibria for a given wage path by characterizing the myopic equilibrium.

Lemma 6
(i) If $\{\tilde{C}^t\}$ *is monotonically decreasing (*$\tilde{C}^{t+1} \leq \tilde{C}^t$*), then* $\{\tilde{x}^t_m, \tilde{x}^t_n, \tilde{C}^t\}$ *is an equilibrium for any* $\delta > 0$*.*
(ii) If $\{\tilde{C}^t\}$ *is monotonically increasing (*$\tilde{C}^{t+1} \geq \tilde{C}^t$*) and* $\sum^k \int_I \hat{a}^t_{(i,k)} \geq \nu$ *for all* t*, then for* $\delta > 0$ any *equilibrium sequence of* $\{C^t\}$ *satisfies* $C^t \geq \tilde{C}^t$ *for all* t*.*

Importantly, Lemma 6 (ii) shows that given $\{w^t\}$ such that the myopic equilibrium is monotonically increasing in $\tilde{C}^t$ and market-clearing, for any $\delta > 0$ the equilibrium sequence $\{C^t\}$ is bounded below by $\{\tilde{C}^t\}$ in *all* dynamic equilibria.

4.2 Dynamic Transition

To formalize the problem of transition introduced at the end of Section 3, we address the question of whether a wage path $\{w^t\}_1^\infty$ exists that shifts the state variable, C^t, from the initial reputation, C^0, to the optimal steady state equilibrium.

Definition 5
A wage path, $\{w^t\}_1^\infty$, transitions *to* $\{w^*, C^*\}$ *if, in all equilibria of the dynamic model, either (1)* $w^t = w^*$ *and* $C^t \geq C^*$ *for some* t*, or (2)* $\lim_{t\to\infty}\{w^t, C^t\} = \{w^*, C^*\}$.

Due to the existence of multiple equilibria in the dynamic setting, there may not exist a wage path that transitions precisely to $C^t = C^*$ in all equilibria. Therefore, we focus on wage paths where C^t is either bounded below by C^* for some t, or approaches C^* as $t \to \infty$.

Given the possibility of multiple equilibria, it will be difficult to provide a direct proof that a wage path either does or does not transition. However, Lemma 6 simplifies the characterization of a transitional wage path by enabling us to focus on the unique myopic equilibrium, $\{\tilde{x}_m^t, \tilde{x}_n^t, \tilde{C}^t\}$. In particular, Lemma 6 ensures that if, given $\{w^t\}$, the myopic equilibrium is monotonically increasing in $\tilde{C}^t$, market-clearing, and transitions to $\{w^*, C^*\}$, then the wage path also transitions for any non-myopic equilibrium. Therefore, we begin the analysis by characterizing the interaction of wages and $\tilde{C}^t$:

Lemma 7 (Crowding out/in motivation)
If $v(C^{t-1}) \leq 0$*, then* $\tilde{C}_1^t$ *is weakly increasing in* w^t.
If $v(C^{t-1}) > 0$*, then* $\tilde{C}_1^t$ *is weakly decreasing in* w^t.

Lemma 7 states that the direction of the effect of current-period wages on the reputation the young cohort depends only on whether the reputation payoff the motivated type receives from employment in institution A is positive or negative: if the reputation payoff is positive, then higher wages crowd out motivated agents; if the reputation payoff is negative, then higher wages *crowd in* motivated agents. The proof follows from the linearity of utility with respect to wages. Intuitively, since the distance between $\tilde{x}_m^t$ and $\tilde{x}_n^t$ is fixed by $v(C^{t-1})$, a wage increase effectively adds a mass of young workers to institution A who have an average motivation of λ, implying that higher wages result in a reputation closer to λ.[17]

Lemma 7 also provides insight regarding potential transition paths between an initial steady-state with low reputation, $\{\bar{w}_0^A, \bar{C}_0\}$, and the optimal steady-state, $\{w^*, C^*\}$:

Corollary 2 (Non-Monotonic Transition)
Given $v(C^0) < 0$*, no monotonic wage path exists that transitions to* $\{w^*, C^*\}$.

Corollary 2 follows from Lemma 6 (i) and the comparative statics outlined in Lemma 7, and shows that given a low initial reputation, a transition cannot be achieved by a wage

[17] This will not always be true locally for non-uniform distributions of $x_{(i,k)}$; however, the result holds more generally for changes in the wage that are large enough, implying that a transition similar to the path illustrated in Proposition 4 below will exist for non-uniform ability distributions.

path that simply decreases w to w^*.[18] Therefore, if a transition path exists from an initial low-motivation steady-state, it must be non-monotonic.

The next result characterizes a non-monotonic wage path that enables a transition to the optimal steady-state equilibrium. To keep the exposition as simple as possible, the following proposition only applies to market-clearing optimal steady states – we extend the result to include non-market-clearing optimal steady states in the Appendix.

Proposition 4 (Existence of Transition)
Given a market-clearing optimal steady state, the following non-monotonic wage path, $\{w^t\}'$, transitions $\{w^t, \tilde{C}^t\}$ from any $\{w^0, C^0\}$ to $\{w^*, C^*\}$:

1. *w^1 and w^2 solve $w^t + v(\tilde{C}^{t-1}) = \overline{x}$.*
2. *w^t for $t > 2$ solves $\int_I \hat{a}_{i,t} = \nu/2$.*

To build intuition for our main result, we first provide a general intuition that links the dynamics of the OLG model to the equilibria of the static model. We then directly illustrate how this wage path results in a transition by solving for the market-clearing wage and the corresponding $\tilde{C}^t$ in each period.

A general intuition for the non-monotonic transition of the myopic equilibrium comes from the link between the dynamics of the OLG model and the equilibria of the static model. In particular, given w^t and C^{t-1}, take x'_m such that $C(x'_m) = C^{t-1}$. If $u_m(w, C(x'_m))$ is greater (lesser) than x'_m, then $\tilde{C}^t_1$ is greater (lesser) than C^{t-1}, since $u_m(w^t, C(x'_m)) = \tilde{x}^t_{m,1}$ in the myopic equilibrium. Therefore, as illustrated in left-hand-side graph of Figure 2, given a constant wage $\bar{w}$ and an initial reputation C^{t-1}, the myopic equilibrium will converge to either the low-motivation or the high motivation steady state that corresponds to a stable static equilibrium, depending on where x'_m is located on the graph.[19]

However, as seen in the right-hand-side graph of Figure 2, if $\bar{w}$ is high enough so that there is a unique steady state, then the myopic equilibrium will converge towards this high-motivation point. This implies that given $w^t = \bar{w}$ for a sufficient number of periods, the myopic equilibrium will reach a point such that $\tilde{C}^{t-1} > \lambda$. Moreover, given $C^{t-1} > \lambda$, x'_m will always be in the "basin of attraction" of the high-motivation steady state (see the left-hand-side graph of Figure 2). This allows a transition to the optimal steady-state equilibrium since, even if w^t is decreased to a point where there exist multiple steady-states, the myopic equilibrium will always converge to the high-motivation steady state.

Next, we illustrate how the wage path specified in Proposition 4 results in a transition by solving for the market-clearing wage and the corresponding $\tilde{C}^t$ in each period.[20]

[18] Note that $v(C^0) < 0$ implies $C^0 < \lambda$, but the reverse need not be true; a transition can be achieved with a monotonic wage path with $C^0 < \lambda$ and $v(C^0) > 0$ since, by Lemma 5, C^1 will be greater than λ.

[19] The myopic equilibrium will converge monotonically to the steady state if each period equilibrium is market-clearing, since the size of each cohort in A remains constant in each period.

[20] Importantly, since the wage is either at or above the market-clearing level in all periods, the size of the

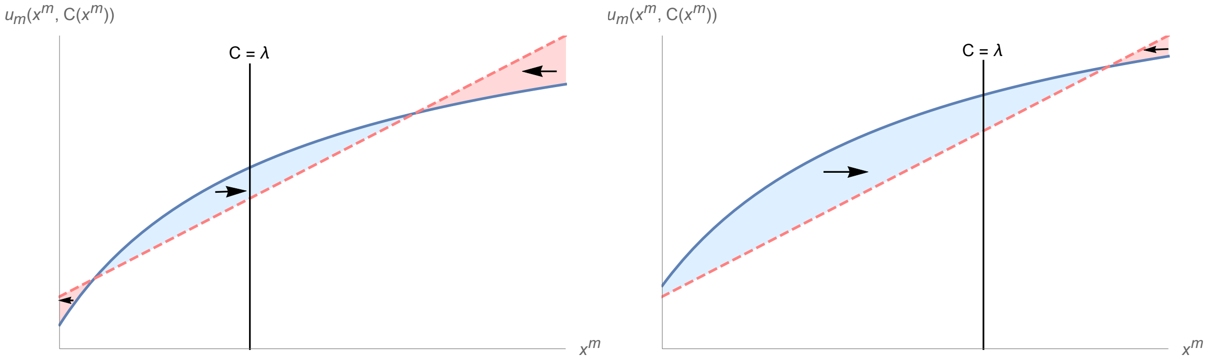

Figure 2: These graphs illustrate the dynamics of $\tilde{C}_1^t$ as a function of C^{t-1} for a given wage: given $u_m(w, C(x'_m))$ greater (lesser) than x'_m, $\tilde{C}_1^t$ is greater (lesser) than C^{t-1}.

Initially, institution A is endowed with low-motivation reputation, $v(\tilde{C}^0) < 0$, and by Lemma 7, A's reputation can only be increased by increasing wages. Taken to the extreme, w^1 and w^2 are set such that all workers prefer employment in institution A, which implies that A's reputation in period 2 will replicate the population average ($\tilde{C}^2 = \lambda$).

In period 3, given that the reputation payoffs are positive, $v(\tilde{C}^2 = \lambda) > 0$, the cutoff level $\tilde{x}_{m,1}^3$ is higher than $\tilde{x}_{n,1}^3$, which implies that $\tilde{C}_1^3 > \lambda$ for the market clearing wage. Moreover, the market clearing wage must be smaller than w^2, since at $w^3 = w^2$, all young agents would set $\hat{a}_{(i,1)} = 1$, which implies that institution A would be over-subscribed. Together, this implies that in period 3, $w^3 < w^2$, and $\tilde{C}^3 > \tilde{C}^2$.

In period 4, the market-clearing wage, w^4 is strictly lower than w^3, since $\tilde{C}^3 > \tilde{C}^2$, again making employment in institution A relatively more attractive for motivated workers in the young cohort. Furthermore, by Lemma 7, $w^4 < w^3$ implies that $\tilde{C}^4 > \tilde{C}^3$. By this same logic, in all future periods the market-clearing wage is decreasing and A's reputation is increasing. Moreover, since the sequences $\{w^t\}$ and $\{\tilde{C}^t\}$ are bounded and monotonic, $\{w^t, \tilde{C}^t\}$ must converge as $n \to \infty$. Lastly, the points of convergence must correspond to $\{w^*, C^*\}$, since this is the unique market-clearing, high-motivation steady state.

The transition outlined above characterizes the general shape of the non-monotonic wage path (also illustrated visually in Figure 3). Starting from a point of low-motivation, w^t must be increased to induce motivated workers to join institution A, hence "purchasing" a higher reputation. Once a sufficiently high reputation has been reached, the process is reversed and public-sector wages are lowered, disproportionately driving non-motivated workers out of the public sector.

While Proposition 4 demonstrates that a transitional wage path exists with myopic agents, Lemma 6 allow us to generalize the result. As seen in Figure 3, given the wage

young and old cohort of agents in A will always be equal ($\pi_1 = \pi_2 = 1/2$), allowing us to disregard cohort effects.

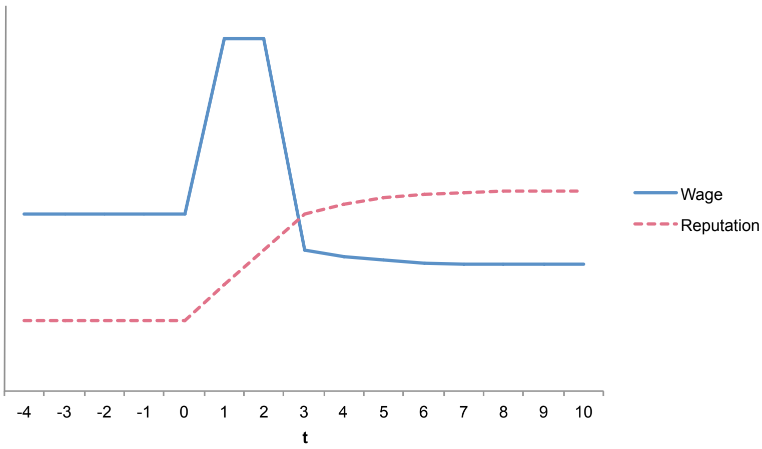

Figure 3: This graph illustrates a wage path that transitions from a low-motivation point to the optimal high-motivation steady-state. Note the initial increase in the wage (solid line) in the mission-oriented institution, followed by a decrease and convergence to the efficiency wage. Reputation (dashed line), however, increases monotonically.

path outlined in Proposition 4, $\{\tilde{C}^t\}$ is monotonically increasing. Therefore, Lemma 6 (ii) applies, which gives the following Corollary:

Corollary 3 (Transition in all equilibria)
For any value of δ and C^0, there exists a wage path $\{w^t\}'$ that transitions to $\{w^*, C^*\}$.

That is, multiple equilibria may exist for any given wage path when motivated agents are forward-looking, since the decision to join A depends on the decision of other motivated agents (through C^t). However, given that the wage path illustrated in Proposition 4 satisfies the conditions of Lemma 6 (ii), the resulting $\{C^t\}$ is bounded below by $\{\tilde{C}^t\}$ *in all equilibria*. This implies that $\{w^t\}'$ transitions to $\{w^*, C^*\}$ *even* when motivated agents are forward looking.

Lastly, while the main focus of this research is on mission-oriented institutions, we consider the case of a "generic" institution, and show that mission is a necessary condition for the existence of a transition in all equilibria.

Proposition 5 (Non-existence of transition to $\{w^*, C^*\}$ with $v(\lambda) = 0$)
If $v(\lambda) = 0$ and $v(C^0) < 0$, then for any $\{w^t\}$ there exists an *equilibrium such that $C^t \leq \lambda$ for all t.*

The intuition behind the contrast of Corollary 3 and Proposition 5 follows from the fact that the existence of a transition path depends on achieving a reputation such that $v(\tilde{C}^t) > 0$ through a wage increase. If this is not possible, then the parameter region where the proportion of motivated workers is decreasing in w^t cannot be reached, and a transitional wage path does not exist. In a mission-oriented institution, we showed that given a high enough wage, $\{\tilde{C}^t\}$ will converge to a point where $v(\tilde{C}^t) > 0$. In a generic institution, however, given a high enough wage, $\{\tilde{C}^t\}$ will converge to $\tilde{C}^t = \lambda$ (from below), which implies that $v(\tilde{C}^t) < v(\lambda) = 0$ for all t. Therefore, it is only in a mission-oriented institution that a reputation such that $v(\tilde{C}^t) > 0$ can be achieved through a wage increase, which enables a transition that is unavailable to generic firms.

4.3 Cost-Minimizing Transition Path

While the transition wage path outlined above relatively quickly reforms institution A's reputation from low to high motivation, it requires a high outlay of wages in the first two periods. Given that many public institutions may face legal and budgetary constraints that limit their ability to raise wages in any given period, in this section we detail the transition path that minimizes the maximum wage bill of institution A.[21] To avoid an open-set problem, we assume that wages can only be increased in multiples of a discrete unit, ϵ, where ϵ is arbitrarily small:

Proposition 6 (Minimum Budget)
The following wage path, $\{w^t\}'$, minimizes the maximum period budget required to transition from C^0 to $\{w^*, C^*\}$:

$$w^t = \begin{cases} \underline{x} + v(0) + \epsilon & \text{for } t \text{ s.t. } v(\tilde{C}^{t-1}) \leq 0 \\ w' \text{ where } w' \text{ solves } \tilde{C}_1^t = C^* & \text{for } t' = \min\{t | v(\tilde{C}^{t-1}) > 0\} \text{ and } t'+1 \\ w^* & \text{for } t > t'+1. \end{cases}$$

The proof of Proposition 6 can be demonstrated using the best-response dynamics of the static model since, as showed above, given a fixed wage the myopic dynamic equilibrium converges to a stable equilibrium of the static model. As illustrated in Figure 2, given $w^t > \underline{x} + v(0)$, the unique static equilibrium is high-motivation, and the myopic equilibrium will converge to this point. In other words, if the wage is set high enough that the motivated worker with the lowest ability wishes to join the public sector, then there is a unique stable equilibrium of the static model at $C' > \lambda$. Therefore, the myopic equilibrium converges to C', and t' such that $v(C^{t'-1}) > 0$ will be reached in finite time.

Proposition 6 also shows that as soon as t' such that $v(C^{t'-1}) > 0$ is reached, then a transition to the efficient steady-state can be achieved in two periods. This transition path implies that institution A will not achieve a maximum profit in periods t' and $t'+1$, since w' will result in institution A being under-demanded in these periods. However, this is a cost-effective strategy, since it achieves a faster transition to the optimal steady-state equilibrium.[22]

[21] The "utility-optimal" transition path is not well-defined, given that multiple equilibria exist for most wage paths that result in a transition.

[22] This strategy of transition may not be robust, since a large drop in salary may disturb the employer-employee relationship by, for example, eroding trust (see Sliwka, 2007 and Ellingsen and Johannesson,

4.4 Extension: Correlation between Ability and Motivation

It is natural to imagine that an agent's ability and their level of motivation may be correlated. For example, due to either selection or socialization, it is often suggested that individuals pursuing a degree in economics are less prosocial and, depending on the mission of their workplace, may therefore be less motivated. Here we consider the case where ability in institution A is heterogenous and correlated with motivation, and show that this extension of the model suggests potentially important insights for transitioning to a high-motivation steady state using selective hiring.

Additionally, the extension we analyze here relaxes the assumption that workers are ex ante identical from A's perspective: since we assume that worker ability is observable, A receives a signal that is correlated a worker's motivation, and can condition their employment contract on this signal. Importantly, we show that while the possibility of conditional contracts may facilitate a transition, the transitional wage path is still non-monotonic.

We amend the baseline model by introducing heterogeneity in $y_{(i,k)}$, (i,k)'s ability in institution A. Specifically, $y_{(i,k)} \in \{y_1, ..., y_p\}$ and, for simplicity, we assume that there is a measure $1/p$ of each ability-type with compact index set I_p. Additionally, a proportion λ_p of each ability-type is motivated (abusing notation, we use a p superscript to refer to variables that are differentiated by ability in institution A); take $\bar{\lambda}$ to be the average level of motivation of the population, $\bar{\lambda} = \sum_p \lambda_p/p$. To introduce correlation between ability and type, we assume that $\lambda^i \neq \lambda^j$ for some $i, j \in \{1, ..., p\}$.

We also allow for a correlation between $y_{(i,k)}$ and $x_{(i,k)}$ by allowing the support of the outside option to condition on p; that is, given $y_{(i,k)} = y_p$, the outside option is uniformly distributed over $[\underline{x}_p, \overline{x}_p]$. Again, to focus on the problem of selecting based on motivation, $y_{(i,k)}$ is publicly observed while the outside option and motivation is unobservable. Analogous to the model above, institution A has a unit demand of $\nu_p < 1/p$ of each ability-type, and sets a uniform wage conditional on $y_{(i,k)}$, w_p^t.

Lastly, take C_p^t to equal the average level of motivation by ability-level:

$$C_p^t = \frac{\sum^k \int_{I_p} m_{(i,k)} a_{(i,k)}}{\sum^k \int_{I_p} a_{(i,k)}},$$

while, as before, C^t equal to the average reputation of institution A.

The following proposition illustrates that, depending on the precise nature of reputation-payoffs, the correlation between ability and motivation can be exploited to transition between a low-motivation point and the high-motivation steady state equilibrium (the argument for existence of a high-motivation steady state equilibrium is analogous to the

2008). Therefore, we have highlighted transition paths that involve market-clearing wages.

baseline model).

Proposition 7 (Transition of Average Reputation)
(i) If motivated agents value ability-contingent reputation, i.e. $u_m^t(w^t, C^{t-1}, a^t_{(i,k)}) = w^t + v(C_p^{t-1})a^t_{(i,k)}$*, then a wage path that transitions from* $v(C^0) < 0$ *to* $\{w^*, C^*\}$ *in all equilibria exists if, and only if,* $v(\lambda_p) > 0$.
(ii) If motivated agents value the average reputation of the institution, i.e. $u_m^t(w^t, C^{t-1}, a^t_{(i,k)}) = w^t + v(C^{t-1})a^t_{(i,k)}$*, then a non-monotonic wage path that transitions from* $v(C^0) < 0$ *to* $\{w^*, C^*\}$ *in all equilibria exists if* $v(\bar{\lambda}) \geq 0$.

Result (i) is a straightforward corollary of Corollary 3 and Proposition 5: if motivated agents value ability-contingent reputation, then each ability category can be treated as its own institution. Result (ii), however, illustrates that if motivated agents value average reputation, then institution A can exploit the correlation between ability and motivation to transition to a high-motivation steady state even if $v(\bar{\lambda}) = 0$. This result follows from the simple intuition that A can manipulate its reputation by disproportionately hiring agents from ability levels with average levels of motivation above that of the population average.

However, a simple strategy of only disproportionately hiring agents high-motivation ability types will not result in a transition. Setting $w_p^1 = w_p^0$ for any non-empty subset of ability types with $\lambda_p > \bar{\lambda}$ and $w_p^1 = 0$ for all others will result in an increase in reputation ($C^1 > C^0$). However, C^1 will still be smaller than $\bar{\lambda}$, since $x^1_{m,p} < x^1_{n,p}$ for each p, given that $v(C^0) < 0$. Moreover, the comparative statics of Lemma 7 hold for each ability-type individually. Therefore, a non-monotonic wage path is still required to ensure a transition to the optimal steady state for any $v(\bar{\lambda}) \geq 0$, which shows that the main result is robust to a setting where institution A receives an imperfect signal of each agent's level of motivation.

The insights from this section may inform a commonly-attempted strategy for reforming an institution by creating an "elite" division within the institution that is staffed by highly-motivated individuals. Logically, this strategy may be successful if recruitment for the elite division targets a group of individuals with a high average motivation, proxied by, say, a certain level educational achievement. However, given a poor institutional reputation, the institution faces an adverse selection problem that may cause this strategy to fail: among the target group, individuals with low motivation are disproportionately attracted to the elite division, since the low collective reputation of the institution dissuades high-motivation individuals from applying. Therefore, this strategy can only be successful if the elite division targets a group of individuals with a high average motivation *and* sets a wage high enough to overcome the adverse selection problem caused by the institution's low collective reputation.

5 Conclusion

While many factors may contribute to the transition to a Weberian public bureaucracy, here we consider the role of selection to the public service in transitioning a public institution from a point of low-motivation to an efficient, high-motivation steady state. In particular, we analyze a model of labor-market sorting with motivated agents who value both mission and collective reputation, and consider the use of wages as a policy tool for reform. We highlight the prediction that the effect of wages on motivation is conditional upon the collective reputation of the mission-oriented institution, which implies that a non-monotonic wage path is needed to transition from a low-motivation to a high-motivation steady-state.

Regarding public-sector reform and wages, it is informative to consider Sweden's transition to a modern, well-functioning bureaucracy, and the path of public-sector wages during this transition. Currently, Swedish bureaucrats enjoy a reputation as belonging to one of the internationally best-regarded systems of public administration.[23] However, in the mid-1800's corruption was endemic to the Swedish system of public administration, where the system of payments to government officials involved "gifts" for services rendered (Rothstein, 2011). It was only after a period of reform, involving a radical transformation of the system of payment that the Swedish bureaucracy evolved into the efficient institution we see today.

Interestingly, while public-sector wages were initially increased to ensure that bureaucrats were not dependent on direct payments from citizens, and to open the profession to a larger class of individuals, public-sector wages fell relative to private-sector wages in "the decades surrounding the Second World War" (Granholm, 2013 pp. 101, translated from Swedish). Moreover, Sundell (2013) provides evidence that nepotism in the Swedish public service decreased continually through this period, and argues that this is indicative of a continuous transition to a well-functioning public sector, a pattern that fits the transition path we outline here.

While the example of Sweden is suggestive, an appropriate test of our model would need to consider the key prediction that the effect of wages on motivation is conditional on the reputation of the public institution. As discussed in the introduction, our results are consistent with the contrasting findings on the effect of higher wages on the average motivation of applicants reported in Dal Bó et al. (2013) and Deserranno (2015). That is, the empirical results can be explained within the context of our model assuming that the hiring institution of the Dal Bó et al. (2013) study, the Mexican federal government, has

[23] Also, in line with the predictions of a public-efficiency wage, these results are achieved with a workforce that is paid less than their peers in the private sector: controlling for observables, de Koning and de Hek, 2013 find an average differential of seven percent among central government workers, and fourteen percent for local government.

a low reputation, while the hiring institution of the Deserranno (2015) study, the well-regarded NGO BRAC, has a high reputation. However, we cannot exclude the possibility that the two studies varied on important dimensions other than the reputation of the hiring institution – a more complete test of the model could be conducted by varying both wages and institutional reputation in a single setting.

We end with some comments on how the mechanism for transition we introduce here can complement other efforts at reform, and by outlining avenues for future research. First, it is important to note that the non-pecuniary motives that we analyze depend on the composition of types in the public institution, rather than the precise level of social output. That is, type-signaling and homophily are independent of the precise behavior of non-motivated and motivated types, as long as there is a difference in behavior between the two types that can be identified through the aggregate behavior of the institution. Therefore, a policy measure such as improved monitoring to target corruption is orthogonal to our mechanism as long as workers update their expectations of each type's behavior.

Additionally, our mechanism is complementary to efforts to change institutional culture by changing institutional norms: If some proportion of workers are conformist (see Bernheim, 1994 and Huck et al., 2012 for models of social norms based on conformity), and hence switch from non-motivated to motivated given some threshold level of aggregate motivation, then increasing the proportion of motivated types in the institution due to selection will precipitate a complementary shift in behavior of the conformist types. This will in turn speed the transition by improving the institution's collective reputation.

Lastly, the model we present here is tailored to address the issue of reform in the mission-oriented sector. However, institutional reputation is also important for targeting talented and motivated individuals in the private sector (see Bhattacharya et al., 2008). And while we address reform of generic institutions here in a partial-equilibrium setting, a relevant extension of our analysis would be to consider reputation in a setting with endogenous mission and competition over motivated workers.

References

Acemoglu, Daron and Matthew Jackson (2015), "History, Expectations, and Leadership in the Evolution of Social Norms." *The Review of Economic Studies*, 82, 423–456.

Acemoglu, Daron and Matthew Jackson (2016), "Social Norms and the Enforcement of Laws." *Journal of the European Economic Association*, Forthcoming.

Akerlof, George and Rachel Kranton (2000), "Economics and Identity." *The Quarterly Journal of Economics*, 115, 715–753.

Akerlof, George and Rachel Kranton (2005), "Identity and the Economics of Organizations." *The Journal of Economic Perspectives*, 19, 9–32.

Aldashev, Gani, Esteban Jaimovich, and Thierry Verdier (2014), "When warm glow burns: Motivational (mis)allocation in the non-profit sector." *Working Paper*.

Ariely, Dan, Anat Bracha, and Stephan Meier (2009), "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially." *The American Economic Review*, 99, 544–555.

Auriol, Emmanuelle and Stefanie Brilon (2014), "Anti-social behavior in profit and non-profit organizations." *Journal of Public Economics*, 117, 149–161.

Banerjee, Ritwik, Tushi Baul, and Tanya Rosenblat (2015), "On self selection of the corrupt into the public sector." *Economics Letters*, 127, 43–46.

Bénabou, Roland and Jean Tirole (2006), "Incentives and Prosocial Behavior." *The American Economic Review*, 96, 1652–1678.

Bénabou, Roland and Jean Tirole (2011), "Identity, Morals, and Taboos: Beliefs as Assets." *The Quarterly Journal of Economics*, 126, 805–855.

Bernheim, B. Douglas (1994), "A theory of conformity." *Journal of Political Economy*, 102, 841–877.

Besley, Timothy and Maitreesh Ghatak (2005), "Competition and Incentives with Motivated Agents." *The American Economic Review*, 95, 616–636.

Besley, Timothy and Maitreesh Ghatak (2016), "Market Incentives and the Evolution of Intrinsic Motivation." *Working Paper*.

Besley, Timothy, Anders Jensen, and Torsten Persson (2014), "Norms, Enforcement, and Tax Evasion." *Working Paper*.

Besley, Timothy and John McLaren (1993), "Taxes and Bribery: The Role of Wage Incentives." *The Economic Journal*, 103, 119–141.

Bhattacharya, C B, Sankar Sen, and Daniel Korschun (2008), "Using Corporate Social Responsibility to Win the War for Talent." *MIT Sloan Management Review*, 49, 37–44.

Bidner, Chris and Patrick Francois (2013), "The Emergence of Political Accountability." *The Quarterly Journal of Economics*, 128, 1397–1448.

Bisin, Alberto and Thierry Verdier (2001), "The Economics of Cultural Transmission and the Dynamics of Preferences." *Journal of Economic Theory*, 97, 298–319.

Caselli, Francesco and Massimo Morelli (2004), "Bad politicians." *Journal of Public Economics*, 88, 759–782.

Cowley, Edd and Sarah Smith (2013), "Motivation and mission in the public sector: evidence from the World Values Survey." *Theory and Decision*, 76, 241–263.

Dal Bó, Ernesto, Frederico Finan, and Martín Rossi (2013), "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service." *Quarterly Journal of Economics*, 128, 1169–1218.

de Koning, Jaap and Paul de Hek (2013), "Wages in the Public Sector: a Comparison Between France, Germany, the Netherlands, Sweden and the United Kingdom." *SEOR Report*.

Delfgaauw, Josse and Robert Dur (2007), "Signaling and screening of workers' motivation." *Journal of Economic Behavior & Organization*, 62, 605–624.

Delfgaauw, Josse and Robert Dur (2008), "Incentives and Workers' Motivation in the Public Sector." *The Economic Journal*, 118, 171–191.

Deserranno, Erika (2015), "Financial incentives as signals: Experimental evidence from the recruitment of Health Workers." *Working Paper*.

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole (1999), "The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies." *Review of Economic Studies*, 66, 199–217.

Di Tella, Rafael and Ernesto Schargrodsky (2003), "The Role of Wages and Auditing during a Crackdown on Corruption in the City of Buenos Aires." *Journal of Law and Economics*, 46, 269–292.

Dixit, Avinash (2002), "Incentive contracts for faith-based organizations to deliver social services." In *Economic theory in a changing world: Policy modeling for growth* (Sajal Lahiri and Pradip Maiti, eds.), Oxford University Press.

Dur, Robert and Robin Zoutenbier (2014), "Intrinsic Motivations of Public Sector Employees: Evidence for Germany." *German Economic Review*, 16, 343–366.

Ellingsen, Tore and Magnus Johannesson (2008), "Pride and Prejudice: The Human Side of Incentive Theory." *The American Economic Review*, 98, 990–1008.

Finan, Frederico, Benjamin Olken, and Rohini Pande (2015), "The Personnel Economics of the State." *NBER Working Paper Series*, 1–48.

Fischer, Paul and Steven Huddart (2008), "Optimal Contracting with Endogenous Social Norms." *The American Economic Review*, 98, 1459–1475.

Francois, Patrick (2000), "'Public service motivation' as an argument for government provision." *Journal of Public Economics*, 78, 275–299.

Francois, Patrick and Michael Vlassopoulos (2008), "Pro-social Motivation and the Delivery of Social Services." *CESifo Economic Studies*, 54, 22–54.

Granholm, Arne (2013), "När Professionerna Blev Byråkrater: Om statlig byråkrati under 200 år och hur traditionella professioners särställning minskat." *Umeå Papers in Economic History.*

Gregg, Paul, Paul Grout, Anita Ratcliffe, Sarah Smith, and Frank Windmeijer (2011), "How important is pro-social behaviour in the delivery of public services?" *Journal of Public Economics*, 95, 758–766.

Gul, Faruk, Hugo Sonnenschein, and Robert Wilson (1986), "Foundations of dynamic monopoly and the coase conjecture." *Journal of Economic Theory*, 39, 155–190.

Handy, Femida and Eliakim Katz (1998), "The Wage Differential between Nonprofit Institutions and Corporations: Getting More by Paying Less?" *Journal of Comparative Economics*, 26, 246–261.

Hanna, Rema and Shing-Yi Wang (2016), "Dishonesty and Selection into Public Service." *AEJ: Economic Policy*, Forthcoming.

Henderson, Rebecca and Eric Van den Steen (2015), "Why Do Firms Have "Purpose"? The Firm's Role as a Carrier of Identity and Reputation." *American Economic Review*, 105, 326–330.

Huck, Steffen, Dorothea Kübler, and Jörgen Weibull (2012), "Social norms and economic incentives in firms." *Journal of Economic Behavior & Organization*, 83, 173–185.

Kosfeld, Michael and Ferdinand von Siemens (2011), "Competition, cooperation, and corporate culture." *The RAND Journal of Economics*, 42, 23–43.

Lazarsfeld, Paul and Robert Merton (1954), "Friendship as a Social Process: A Substantive and Methodological Analysis." In *Freedom and Control in Modern Society* (Morroe Berger, Theodore Abel, and Charles Page, eds.), 18–66, Van Nostrand, New York.

Macchiavello, Rocco (2008), "Public sector motivation and development failures." *Journal of Development Economics.*

Morris, Stephen and Hyun Song Shin (2000), "Rethinking Multiple Equilibria in Macroeconomic Modeling." *NBER Macroeconomics Annual*, 15, 139–182.

Murdock, Kevin (2002), "Intrinsic Motivation and Optimal Incentive Contracts." *The RAND Journal of Economics*, 33, 650–671.

Prendergast, Canice (2007), "The Motivation and Bias of Bureaucrats." *The American Economic Review*, 97, 180–196.

Prendergast, Canice (2008), "Intrinsic Motivation and Incentives." *The American Economic Review*, 98, 201–205.

Rose-Ackerman, Susan (1999), *Corruption and Government: Causes, Consequences, and Reform.* Cambridge University Press.

Rothstein, Bo (2011), "Anti-corruption: the indirect 'big bang' approach." *Review of International Political Economy*, 18, 228–250.

Sliwka, Dirk (2007), "Trust as a Signal of a Social Norm and the Hidden Costs of Incentive Schemes." *The American Economic Review*, 97, 999–1012.

Sundell, Anders (2013), "Nepotism in the Swedish Central Public Administration 1790-1925." *Working Paper.*

The Global Journal (2015), *Top 500 NGOs.* Global Geneva.

Tirole, Jean (1996), "A theory of collective reputations (with applications to the persistence of corruption and to firm quality)." *The Review of Economic Studies*, 63, 1–22.

Transparency International (2014), *Corruption Perceptions Index 2014.*

Treisman, Daniel (2000), "The causes of corruption: a cross-national study." *Journal of Public Economics*, 76, 399–457.

Valasek, Justin (2016), "Reforming an Institutional Culture: A Model of Motivated Agents and Collective Reputation." *WZB Discussion Paper SP II 2015-303r.*

Van Rijckeghem, Caroline and Beatrice Weder (2001), "Bureaucratic corruption and the rate of temptation: do wages in the civil service affect corruption, and by how much?" *Journal of Development Economics*, 65, 307–331.

Wilson, James (1989), *Bureaucracy: What Government Agencies Do and Why They Do It.* Basic Books.

6 Appendix A: Proofs

6.1 Proofs for Section 3

Proof of Lemma 1:
First, for non-motivated types, $u_n(w, C, a_i = 1) = w$. Therefore, non-motivated workers have a best-response to set $\hat{a}_i = 1$ iff $x_i < w$.

Second, for motivated types, by contradiction, assume that there exists an equilibrium with $x_j < x_i$ and $\hat{a}_j = 0$, $\hat{a}_i = 1$. This implies that $x_i < w + v(C)$ and $x_j \geq w + v(C)$, a contradiction. ■

Proof of Lemma 2:
Since $x_n = w$, a high-motivation exists only if an equilibrium cutoff, x_m, exists in the range $(w, \overline{x}]$.

First, if $\overline{x} \leq w + v(C(\overline{x}))$, then the corner, $x_m = \overline{x}$, is an equilibrium, and is high-motivation since $w < \overline{x}$. Moreover, it is the unique high-motivation equilibrium since, given that $u_m(w, C(x_m))$ is concave in x_m and $u_m(w, C(x_m)) \geq x_m$ at both w and $\overline{x}$, there is no crossing of $u_m(w, C(x_m))$ and x_m in $[w, \overline{x})$,

Second, by a similar argument, if $\overline{x} > w + v(C(\overline{x}))$, then there exists a unique crossing of $u_m(w, C(x_m))$ and x_m in $(w, \overline{x})$, since $u_m(w, C(x_m)) > x_m$ at w and $u_m(w, C(x_m)) < x_m$ at $\overline{x}$. ■

Proof of Proposition 1:
Since $x_n = w$, a low-motivation exists only if an equilibrium cutoff, x_m, exists in the range $[\underline{x}, \lambda]$.

[Uniqueness] If $w + v(0) > \underline{x}$ then, by the concavity of $v(C(\cdot))$, there is no crossing of $u_m(w, C(x_m))$ and x_m in $[\underline{x}, \lambda]$ since $u_m(w, C(x_m)) > x_m$ at both $\underline{x}$ and λ. Therefore, by Lemma 2, the high-motivation equilibrium is unique.

If $w \in (\underline{x} - v(1), \underline{x})$, then x_n is equal to $\underline{x}$ in any equilibrium, since the wage is below the outside-option of all workers. However, $x^m = w + v(1) > \underline{x}$ is an equilibrium since $C(x^m) = 1$ for any $x^m > \underline{x}$.

[Multiplicity] If $w \in (\underline{x}, \underline{x} - v(0))$, then a crossing of $u_m(w, C(x_m))$ and x_m exists in $[\underline{x}, \lambda]$, since $u_m(w, C(x_m)) < x_m$ at $\underline{x}$ and $u_m(w, C(x_m)) > x_m$ at λ. Additionally, given that $u_m(w, C(0)) < \underline{x}$, the corner $\{x_m = 0, x_n = w, C = 0\}$ is also an equilibrium. ■

Proof of Lemma 3:
(i) By Lemma 2, a unique high-motivation exists for all values of $w \in (\underline{x} - v(1), \overline{x})$. Moreover, in the high-motivation equilibrium, $\int_I a_i \to 0$ as $w^+ \to \underline{x} - v(1)$ and $\int_I a_i \to 1$ as $w^- \to \overline{x}$. Therefore, since both equilibrium cutoff values, x_n and x_m, are continuous in w, a high-motivation equilibrium with $\int_I a_i = \nu$ exists for some value of w in $(\underline{x} - v(1), \overline{x})$.

(ii) For any $w \in (\underline{x}, \underline{x} - v(0))$, a low-motivation equilibrium exists with $x_m = \underline{x}$, since

$w+v(0) < \underline{x}$. Since only non-motivated individuals select into A in this equilibrium, $\int_I a_i$ is equal to $(1-\lambda)(w-\underline{x})$. Therefore, given that at $w' = \underline{x} - v(0)$, $(1-\lambda)(w'-\underline{x}) = (1-\lambda)|v(0)| > \nu$, a low motivation equilibrium exists with $x_m = \underline{x}$ and $\int_I a_i = \nu$. ■

Proof of Proposition 2:

First, we show that the optimal equilibrium cannot be a low-motivation equilibrium. By contradiction, assume that the optimal equilibrium, $\{w^*, C^*\}$, is low-motivation and take $a^* = \int_I \hat{a}_i$ for this equilibrium. By the proof of Lemma 3 (i) there exists a high-motivation equilibrium, $\{w, C^h\}$, with $\int_I \hat{a}_i = a^*$. Since both $\{w^*, C^*\}$ and $\{w, C^h\}$ have the same number of applicants to A, for $\{w^*, C^*\}$ to maximize social profit, it must be the case that $C^* > C^h$ or $w^* < w$. Since $C^* < \lambda < C^h$, w^* must be smaller than w. However, since $\int_I \hat{a}_i = a^*$ and $C^* < C^h$, x_n must be smaller than x_n^*, which implies that $w < w^*$. Therefore, $\{w^*, C^*\}$ must be a high-motivation equilibrium.

Second, to show that w^* is lesser or equal to the wage of a market-clearing high-motivation equilibrium, we utilize the comparative statics of C in the unique high-motivation equilibrium with respect to w. Note that if $\partial C(x_m, w)/\partial w \leq 0$, then $w^* \leq w^h$, since all high-motivation equilibria with $w > w^h$ will have a lower reputation than the market-clearing equilibrium and hence a lower social profit.

Take w' such that the corresponding high-motivation equilibrium $\{x'_m, x'_n, C'\}$ is interior, and consider a wage increase of Δw to w'', with a corresponding high-motivation equilibrium $\{x''_m, x''_n, C''\}$. First, note that $x''_n - x'_n = \Delta x_n = \Delta w$. Next, take $\check{x}_m = x'_m + \Delta w$; since $\check{x}_m - x'_m = \Delta x_m = \Delta x_n$, it must be the case that $C(\check{x}_m, w'') < C(x'_m, w')$. Therefore, $w'' + C(\check{x}_m, w'') < \Delta w + w' + C(x'_m, w') = \Delta x_m + x'_m = \check{x}_m$, which shows that x''_m must be lower than $\check{x}_m$, since $w'' + C(x_m, w'')$ intersects x_m from above in the high-motivation equilibrium. Lastly, since $x''_m < \check{x}_m$, $C(x''_m, w' + \Delta w) < C(\check{x}_m, w' + \Delta w) < C(x'_m, w')$ and, taking the limit as $\Delta w \to 0$, this shows that $\partial C(x_m, w)/\partial w \leq 0$. ■

Lastly, we show that institution A maximizes social profit with a uniform wage. Abusing notation, for this result we use I to refer to the set of workers. For simplicity, we state and prove the result for binary partition of I – the result, however, easily extends to an arbitrary number of partitions. Assume institution A partitions I into I^1 and I^2 with $|I^1| = \theta^1$ and $|I^2| = \theta^2$. Since there is a continuum workers and all workers are identical from A's perspective, workers are effectively randomly allocated to I^1 or I^2 with, respectively, probability θ^1 and θ^2. Therefore, both I^1 and I^2 will have a proportion of motivated workers equal to λ, and $x_i \sim U[\underline{x}, \overline{x}]$. Take w^j to be the wage A offers to the subset of workers I^j.

Result 1 (Uniform wage)

Given any partition, $I^1 \cup I^2 = I$, wages $\{w^1, w^2\}$, and resulting equilibrium $\{x^1_m, x^2_m, x^1_n, x^2_n, C\}$, there exists a uniform wage, w', with a corresponding equilibrium, $\{x'_m, x'_n, C'\}$, such that $C' = C$ and $\int_I \hat{a}'_i = \int_I \hat{a}_i$. Moreover, $w' \int_I \hat{a}'_i < w^1 \int_{I^1} \hat{a}_i + w^2 \int_{I^2} \hat{a}_i$.

This shows that for any equilibrium with a non-uniform wage, there exists an equivalent equilibrium with a uniform wage that has higher social profit than the partitioned equilibrium (due to a lower total wage bill). Importantly, this implies, by transitivity, that $\{w^*, C^*\}$ is not dominated by any partitioned equilibrium, since $\{w^*, C^*\}$ is the optimal uniform-wage equilibrium.

Proof of Result 1:
We prove the existence of a corresponding equilibrium by construction. Take w' such that $\int_I \hat{a}'_i(1-m_i) = \int_{I^1} \hat{a}_i(1-m_i) + \int_{I^2} \hat{a}_i(1-m_i)$. Such a wage exists since x'_n, and hence $\int_I \hat{a}'_i(1-m_i)$, is unique and continuous in w. It follows that:

$$\begin{aligned}(1-\lambda)(\theta^1(x^1_n - \underline{x}) + \theta^2(x^2_n - \underline{x})) &= (1-\lambda)(x'_n - \underline{x}) \text{ or,}\\ \theta^1(x^1_n - \underline{x}) + \theta^2(x^2_n - \underline{x}) &= x'_n - \underline{x}. \end{aligned} \tag{3}$$

Next, take $x'_m = x'_n + v(C)$, where C is the reputation in the partitioned equilibrium. Note that $\{x'_m, x'_n, C\}$ is an equilibrium if $C'(w', x'_m) = C$, which is the case if:

$$\lambda(\theta^1(x^1_m - \underline{x}) + \theta^2(x^2_m - \underline{x})) = \lambda(x'_m - \underline{x}). \tag{4}$$

To see that this is the case, note that in the partitioned equilibrium, $x^1_m - x^1_n = x^2_m - x^2_n = v(C)$. Therefore, the left-hand side of expression 4 can be rewritten as:

$$\begin{aligned}&\lambda(\theta^1(x^1_n + v(C) - \underline{x}) + \theta^2(x^2_n + v(C) - \underline{x})) =\\ &\qquad \lambda([\theta^1(x^1_n - \underline{x}) + \theta^2(x^2_n - \underline{x})] + v(C)),\end{aligned}$$

and as shown above, the term in brackets is equal to $(x'_n - \underline{x})$, which gives:

$$\lambda(x'_n + v(C) - \underline{x}) = \lambda(x'_m - \underline{x}).$$

This shows that expression 4 holds, and a uniform wage exists with a corresponding equilibrium, $\{x'_m, x'_n, C'\}$, such that $C' = C$ and $\int_I \hat{a}'_i = \int_I \hat{a}_i$.

Next we show that the total wage bill is lower under the uniform wage. By contradiction, assume $w' \int_I \hat{a}'_i \geq w^1 \int_{I^1} \hat{a}_i + w^2 \int_{I^2} \hat{a}_i$. We assume WOLOG that $w^1 > w^2$. Since $\int_I \hat{a}'_i(1-m_i) = \int_{I^1} \hat{a}_i(1-m_i) + \int_{I^2} \hat{a}_i(1-m_i)$, it follows that $w^1 > w' > w^2$. Therefore, relative to w', the wage bill increase for I^1 must be lower than the wage bill decrease for I^2:

$$\begin{aligned}&(w^1 - w')\theta^1[(1-\lambda)(w^1 - \underline{\mathrm{x}}) + \lambda(w^1 + v(C) - \underline{\mathrm{x}})] \leq\\ &(w' - w^2)\theta^2[(1-\lambda)(w^2 - \underline{\mathrm{x}}) + \lambda(w^2 + v(C) - \underline{\mathrm{x}})]\end{aligned}$$

However, since $x_n = w$ and $\theta^1 + \theta^2 = 1$, expression 3 simplifies to:

$$\theta^1 w^1 + \theta^2 w^2 = w' \Rightarrow (w^1 - w')\theta^1 = (w' - w^2)\theta^2,$$

which implies that the inequality above reduces to:

$$[(1-\lambda)(w^1 - \underline{x}) + \lambda(w^1 + v(C) - \underline{x})] \leq [(1-\lambda)(w^2 - \underline{x}) + \lambda(w^2 + v(C) - \underline{x})],$$

which is a contradiction since $w^1 > w^2$. ■

6.2 Proofs for Section 4.1

Proof of Lemma 4:
Given the constant wage in institution A over both periods of life, a non-motivated worker with $x^t_{(i,k)} = w^{t+1-k}$ will be indifferent between employment in A and their outside option, while workers with $x^t_{(i,k)} < (>)\ w^{t+1-k}$ will strictly prefer employment in A (the outside option) in both periods of life.

Second, consider a motivated worker with $k = 2$. Since $u^t(w^{t-1}, C^{t-1}, a_{(i,2)} = 1)$ is fixed for period t, $u^t(w^{t-1}, C^{t-1}, a_{(i,2)} = 1)$ and $x_{(i,2)}$ satisfy single crossing. The same argument holds for a motivated worker with $k = 1$, since $U^t(m_{(i,1)}, x_{(i,1)}, a^t_{(i,1)} = 1)$ is fixed *given* w^t, C^{t-1} and C^t. ■

Proof of Proposition 3:
First note that if $w^t = \bar{w}$ and $C^t = \bar{C}$ for all t, then $x^t_{m,1} = x^t_{m,2}$ and $x^t_{n,1} = x^t_{n,2}$ since the period utility for both types is constant. Additionally, the constant period utility implies that in a steady-state, the option value of employment in A for the $k = 1$ cutoff type is equal to zero; i.e. $u^{t+1}(w^t, \bar{C}^t, m_{(i,1)}, x^t_{m,1}) = x^t_{m,1}$. The proof of Proposition 3 then follows from an equivalence of the conditions for a steady-state equilibrium and the equilibrium conditions for a static equilibrium, given the constant collective reputation in the steady-state.

That is, given that the cutoff types are indifferent between $\hat{a}^t_{(i,k)} = 0, 1$ in each period of life, $\{\bar{w}, \bar{x}_m, \bar{x}_n, \bar{C}\}$ define a steady-state equilibrium if and only if the following conditions are satisfied:

$$\begin{aligned}
&\bar{x}_m = \bar{w} + v(\bar{C}),\\
&\bar{x}_n = \bar{w},\\
&\bar{C} = \phi^t_1 C^t_1(\bar{x}_m, \bar{x}_n) + \phi^t_2 C^t_2(\bar{x}_m, \bar{x}_n) = \frac{\lambda(\bar{x}_m - \underline{x})}{(1-\lambda)(\bar{x}_n - \underline{x}) + \lambda(\bar{x}_m - \underline{x})},
\end{aligned}$$

which are equivalent to the conditions for a static equilibrium. ■

Proof of Corollary 1:

Since the period payoffs in a steady-state equilibrium, where $x^t_{m,1} = x^t_{m,2} = \bar{x}_m$, are equivalent to the payoffs of the corresponding static equilibrium, $\Pi^A = \sum(1-\tau)^{t-1}\bar{\pi}^A$ is maximized at the steady-state that maximizes $\bar{\pi}^A$. ■

Proof of Lemma 5:

The proof of uniqueness follows from the fact that workers maximize their objective using the simple decision rule:

$$\hat{a}^t_{(i,k)} = 1 \text{ } \textit{iff} \text{ } w^{t+1-k} + m_{(i,k)} v(C^{t-1}) \geq x_{(i,k)},$$

which implies a unique value for both $\{x^t_{m,k}\}$, $\{x^t_{n,k}\}$ given C^{t-1}. In turn, since the value of C^t is unique given $x^t_{m,k}$, $x^t_{n,k}$, the equilibrium sequence of $\{C^t\}$ is also unique given the initial reputation and workforce, C^0, $\{a^0_{(i,1)}\}^I$. ■

Proof of Lemma 6:

For (i), note that $\{\tilde{x}^t_m, \tilde{x}^t_n, \tilde{C}^t\}$ is a dynamic equilibrium for $\delta > 0$ if $u^{t+1}(w^t, \tilde{C}^t, m_{(i,1)} = 1, \tilde{x}^t_{m,1}) = \tilde{x}^t_{m,1}$ for all t, since this implies that there is no "option value" of employment in A for the cutoff type. This allows us to prove the result by induction: take $\{w^t\}$ such that $\tilde{C}^t \leq \tilde{C}^{t-1}$ for all t. Given $C^{t-1} = \tilde{C}^{t-1}$, take $x^t_{m,1} = \tilde{x}^t_{m,1}$, which gives $C^t = \tilde{C}^t$, since $x^t_{m,2} = \tilde{x}^t_{m,2}$ in all period-equilibria. Therefore $C^t \leq C^{t-1}$, which implies that $u^{t+1}(w^t, \tilde{C}^t, m_{(i,1)} = 1, x^t_{m,1}) = x^t_{m,1}$. This shows that $\{x^t_{m,1}, x^t_{m,2}\} = \{\tilde{x}^t_{m,1}, \tilde{x}^t_{m,2}\}$ is an equilibrium given $C^{t-1} = \tilde{C}^{t-1}$. Since $C^0 = \tilde{C}^0$, this proves the result.

For (ii), we prove the result by contradiction. That is, given $\{w^t\}$, assume there exists an equilibrium, $\{x^t_m, x^t_n, C^t\}$ with $C^t < \tilde{C}^t$ for some t. Take t' to equal the first period in which $C^t < \tilde{C}^t$ (note that $\min\{t|C^t < \tilde{C}^t\} > 0$, since $C^0 = \tilde{C}^0$). Given that $C^{t'-1}$ must be greater or equal to $\tilde{C}^{t'-1}$, it follows that the utility the motivated type receives from employment in A in period t' is weakly higher relative to the myopic equilibrium; that is, $u^{t'}(w^{t+1-k}, C^{t-1}, a^t_{(i,k)} = 1) \geq u^{t'}(w^{t+1-k}, \tilde{C}^{t-1}, a^t_{(i,k)} = 1)$ for $k = 1, 2$. Therefore, given that $k = 1$ retains their outside option in $t' + 1$ and $k = 2$ maximizes their period payoff, it must be the case that the motivated cutoff types in period t' are weakly greater in $\{x^t_m, x^t_n, C^t\}$ relative to $\{\tilde{x}^t_m, \tilde{x}^t_n, \tilde{C}^t\}$.

The fact that $x^{t'}_{m,k} \geq \tilde{x}^{t'}_{m,k}$ for $k = 1, 2$, however, does not contradict $C^{t'} < \tilde{C}^{t'}$: as long as $\phi^{t'}_k \neq \tilde{\phi}^{t'}_k$, it is possible for $C^{t'} = \phi^{t'}_1 C^{t'}_1 + \phi^{t'}_2 C^{t'}_2$ to be strictly smaller than $\tilde{C}^t = \tilde{\phi}^{t'}_1 \tilde{C}^{t'}_1 + \tilde{\phi}^{t'}_2 \tilde{C}^{t'}_2$ even if $C^{t'}_k \geq \tilde{C}^{t'}_k$ for $k = 1, 2$.

Therefore, to complete the proof, we show that $\phi^{t'}_k = \tilde{\phi}^{t'}_k = 1/2$. This result stems from the fact that A is (weakly) over-demanded in $\{w^t, \tilde{x}^t_m, \tilde{x}^t_n, \tilde{C}^t\}$. Therefore, since $\int_I a^t_{(i,1)} \geq \nu/2$ in each period, $\tilde{\phi}^t_k = 1/2$ for all t. Moreover, since $x^t_{m,k} \geq \tilde{x}^t_{m,k}$ for $t \leq t'$ (by the same argument as above), A must also be over-demanded in $\{x^t_m, x^t_n, C^t\}$ all $t \leq t'$. This shows that $\phi^{t'}_k = 1/2$, which completes the proof. ■

Proof of Lemma 7:
First, we give the expression for $\tilde{C}_1^t$ as a function of the cutoff types:

$$\tilde{C}_1^t = \frac{\int_I \hat{a}^t_{(i,1)} m_{(i,1)}}{\int_I \hat{a}^t_{(i,1)}} = \frac{\lambda(\tilde{x}^t_{m,1} - \underline{x})}{(1-\lambda)(\tilde{x}^t_{n,1} - \underline{x}) + \lambda(\tilde{x}^t_{m,1} - \underline{x})}$$

Due to the quasi-linearity of both types' utility with respect to the wage, $\partial \tilde{x}^t_{m,1}(w^t, C^{t-1})/\partial w^t = \partial \tilde{x}^t_{n,1}(w^t)/\partial w^t = 1$ for interior values, which implies that:

$$\partial \tilde{C}_1^t / \partial w^t = \frac{\lambda((1-\lambda)(\tilde{x}^t_{n,1}(w^t) - \underline{x}) + \lambda(\tilde{x}^t_{m,1}(w^t, C^{t-1}) - \underline{x})) - \lambda(\tilde{x}^t_{m,1}(w^t) - \underline{x})}{((1-\lambda)(\tilde{x}^t_{n,1}(w^t) - \underline{x}) + \lambda(\tilde{x}^t_{m,1}(w^t, C^{t-1}) - \underline{x}))^2}$$

This expression if negative iff:

$$(1-\lambda)\tilde{x}^t_{n,1}(w^t) + \lambda \tilde{x}^t_{m,1}(w^t, C^{t-1}) < \tilde{x}^t_{m,1}(w^t),$$

which is true iff $\tilde{x}^t_{m,1}(w^t, C^{t-1}) > \tilde{x}^t_{n,1}(w^t)$.

Next, note that the relationship between $\tilde{x}^t_{m,1}(w^t, C^{t-1})$ and $\tilde{x}^t_{n,1}(w^t)$ depends only on the sign of $v(C^{t-1})$, since motivated types' utility is separable with regard to wage and reputation. In particular:

$$\tilde{x}^t_{m,1}(w^t, C^{t-1}) = w^t + v(C^{t-1}) \lesseqgtr w^t = \tilde{x}^t_{n,1}(w^t, C^{t-1}) \text{ iff } v(C^{t-1}) \lesseqgtr 0.$$

Lastly, note that the same relationship holds when one of the two cutoffs is non-interior, and when both are non-interior, $\partial \tilde{C}_1^t / \partial w^t = 0$. ■

Proof of Corollary 2:
First, note that a monotonically increasing wage path cannot transition to $\{w^*, C^*\}$ since, by Lemma 3, $w^0 > w^*$.

Second, for a monotonically decreasing wage path, $\tilde{C}_1^1 \leq C^0$ by Lemma 7, since $v(C^0) < 0$ and $w^1 \leq w^0$. Therefore, $\tilde{C}^1 \leq \tilde{C}^0$, which implies that $v(\tilde{C}^1) \leq v(\tilde{C}^0) < 0$, and by the same argument, $\tilde{C}_1^2 \leq \tilde{C}_1^1$. Again, this shows that $\tilde{C}^2 \leq \tilde{C}^1$, and by induction $\tilde{C}^{t+1} \leq \tilde{C}^t$ for all t. Therefore, Lemma 6 (i) applies and $\{\tilde{x}^t_m, \tilde{x}^t_n, \tilde{C}^t\}$ is a dynamic equilibrium for any $\delta > 0$. Lastly, the result follows since $\tilde{C}^t \leq C^0 < C^*$ for all t. ■

Before the proof, we state an more general version of Proposition 4.

Proposition 4' (Existence of Transition)
Given $\delta = 0$, there exists a wage path that transitions to $\{w^*, C^*\}$ for any C^0.

Proof of Proposition 4':
The proposition is proved by construction in the main text for the case of a market-clearing optimal steady state. If $\{w^*, C^*\}$ is not market-clearing, then the following wage

path transitions to $\{w^*, C^*\}$:

1. w^1 and w^2 solve $w^t + v(\tilde{C}^{t-1}) = \overline{x}$.
2. w^3 and w^4 solve $\tilde{C}_1^t = C^*$.
3. $w^5 = w^*$.

Given that all workers set $\hat{a}_{(i,k)} = 1$ in periods 1 and 2, $\tilde{C}^2 = \lambda$. For period 3, note that with $v(\tilde{C}^2 = \lambda) > 0$, $\tilde{C}_1^t(w^t)$ is a continuous function with a range of $[\lambda, 1]$. Therefore, there exists a wage such that $\tilde{C}_1^t = C^*$. The same holds for period 4. Finally, since $\tilde{C}^2 < \tilde{C}^3 < \tilde{C}^4$, all workers matched with A in their first period of life remain in A in their second period of life. Therefore, $\tilde{C}^4 = \tilde{\phi}_1^4\tilde{C}_1^4 + \tilde{\phi}_2^4\tilde{C}_2^4 = \tilde{\phi}_1^4 C^* + \tilde{\phi}_2^4 C^* = C^*$, which proves that a transition is achieved in period 5. ■

Proof of Corollary 3:

First, if $\{w^*, C^*\}$ is market-clearing, then the result follows directly from Proposition 4 and Lemma 6, since the myopic equilibrium is market-clearing and increasing in $\{\tilde{C}^t\}$.

If $\{w^*, C^*\}$ is not market-clearing, however, then Lemma 6 no longer applies. However, given $\delta > 0$, the wage path outlined in the above proof of Proposition 4' still transitions in any equilibrium. To see this, first note that the equilibrium is unique in periods 1 and 2, since $u^t(w^{t+1-k}, C^{t-1}, m_{(i,k)}) \geq x_{(i,k)}^t$ for all agents, which implies that $C^2 = \lambda$ in all equilibria. Moreover, given $C^{t-1} = \tilde{C}^{t-1}$, $C_1^t \geq \tilde{C}_1^t$, since workers retain their outside option in their second period of life. Therefore, $C^3 \geq \tilde{C}^3$, which in turn implies $C_k^4 \geq \tilde{C}_k^4$ for $k = 1, 2$ in all equilibria, which proves the result. ■

Proof of Proposition 5:

We prove the result by construction. Take an arbitrary equilibrium, $\{x_m^t, x_n^t, C^t\}$. For a transition to occur, it must be true that $C^{t-1} \leq \lambda$ and $C^t > \lambda$ for some t; take $\hat{t}$ to be the minimum t such that $C^t > \lambda$. This implies that either $C_1^{\hat{t}} > \lambda$ or $C_2^{\hat{t}} > \lambda$ (or both). However, since $C^{\hat{t}-1} \leq \lambda$ implies that $v(C^{\hat{t}-1}) \leq 0$, it follows that $C_2^{\hat{t}} \leq \lambda$. Therefore, $C_1^{\hat{t}} > \lambda$.

Next, we show that an equilibrium, $\{x_m^{t'}, x_n^{t'}, C^{t'}\}$, exists that is identical to $\{x_m^t, x_n^t, C^t\}$ for $t < \hat{t}$, and where $C_1^{\hat{t}'} \leq \lambda$. Take $x_{m,1}^{\hat{t}'} = x_{n,1}^{\hat{t}'} = w^{\hat{t}}$. At this cutoff, $C_1^{\hat{t}'} = \lambda$, which implies that $C^{\hat{t}'} \leq \lambda$. Therefore, the dynamic utility of the cutoff type, $U(w^{\hat{t}}, \{C^{t'}\}, m_{(i,1)}^{\hat{t}}, x_{m,1}^{\hat{t}'}, a_{(i,1)}^{\hat{t}} = 1) = u^t(w^{\hat{t}'}, C^{\hat{t}'-1}, m_{(i,1)}) + \delta u^{t+1}(w^{\hat{t}'}, C^{\hat{t}'}, m_i, x_{m,1}^{\hat{t}'})$ is smaller or equal to $(1+\delta)x_m^{\hat{t}'}$. Moreover, given $x_{m,1}^{\hat{t}} = \tilde{x}_{m,1}^{\hat{t}}$, $U(w^{\hat{t}}, \{C^{t'}\}, m_{(i,1)}^{\hat{t}}, x_{m,1}^{\hat{t}'}, a_{(i,1)}^{\hat{t}} = 1) \geq (1+\delta)\tilde{x}_{m,1}^{\hat{t}}$ by definition. This shows that, by continuity, an equilibrium value of $x_{m,1}^{\hat{t}'}$ exists in the interval $[\tilde{x}^{\hat{t}}, w^{\hat{t}}]$. Therefore $C^{\hat{t}'} \leq \lambda$ in this equilibrium, since $x_{m,1}^{\hat{t}'} \leq w^{\hat{t}} = x_{n,1}^{\hat{t}'}$.

This shows that for any equilibrium that transitions to a high-motivation reputation by period t, there exists an equivalent equilibrium that does not transition ($C^t \leq \lambda$). Since this argument can be applied iteratively, it shows that for any wage path $\{w^t\}$, an equilibrium exists where $C^t \leq \lambda$ for all t. ■

Proof of Proposition 6:

Given that the proof of Proposition 4' establishes that $\{w^t\}'$ transitions once a point with $v(C^{t-1}) > 0$ is reached, it suffices to show that setting a constant wage of $\bar{w} = \underline{x} + v(0) + \epsilon$ will cause a transition to a point such that $v(C^{t-1}) > 0$ in all equilibria. Also, note that by Propositions 1 and 3, given $\bar{w}$ there is a unique steady state with a reputation $C^h > \lambda$.

Similar to the earlier proofs, we will utilize the myopic equilibrium, and begin by proving the following result:

Result 2
Given $w^t = \bar{w}$ and $C^{t-1} < C^h$, $\tilde{C}_1^t > C^{t-1}$.

By contradiction, assume that $\tilde{C}_1^t \leq C^{t-1}$. Take $x'_{m,1} = x_{m,1}$ such that $x_{m,1}$ solves $\tilde{C}_1^t(x_{m,1}) = C^{t-1}$; it follows that $\tilde{x}_{m,1}^t = u_m^t(\bar{w}, C^{t-1}) \leq x'_{m,1}$. This, however, is a contradiction since, given $u_m(\bar{w}, C(\underline{x})) > \underline{x}$, $u_m(\bar{w}, C(x_m)) > x_m$ for all $C(x_m) < C^h$. This proves Result 2.

Next, we show by induction that $\{\tilde{C}^t\}_0^{t'} \subset \{\tilde{C}^t\}_0^\infty$ is strictly increasing for any set such that $\tilde{C}^t < C^h$ for all t. Since $C_1^0 = C^0$ and $C^0 < C^h$, $\tilde{C}_1^1 > C_1^0$ by Result 2, which implies that $\tilde{C}^1 > C^0$. By the same argument $\tilde{C}_1^t > \tilde{C}^{t-1}$ for any $\tilde{C}_1^t \in \{\tilde{C}_1^t\}_0^{t'}$, which also implies that $\tilde{C}^t > \tilde{C}^{t-1}$, since $\tilde{C}_2^t \geq \tilde{C}^{t-1}$.

This proves that either $\tilde{C}^{t-1} \geq C^h$ is reached for some t (in which case $v(\tilde{C}^{t-1}) > 0$), or that $\{\tilde{C}^t\}_0^\infty$ is strictly increasing. In the latter case, note that $\lim_{t\to\infty} \tilde{C}^t$ must equal C^h, since $\{\tilde{C}^t\}$ cannot converge to any other point given that C^h is the unique steady state, and $\{\tilde{C}^t\}$ must converge since it is monotonic and bounded. Lastly, since $v(C^h) - \epsilon > 0$ for some $\epsilon > 0$, if $\{\tilde{C}^t\}$ converges to $v(C^h)$, then $v(\tilde{C}^t) > 0$ for some finite t.

We have shown that given $w^t = \bar{w}$, a subset $\{\tilde{C}^t\}_0^{t'} \subset \{\tilde{C}^t\}_0^\infty$ exists where $v(\tilde{C}^{t'}) > 0$, and $\tilde{C}^t$ is monotonically increasing over this subset. Moreover, since $\bar{w} > w^0$ and $\tilde{C}^t > C^0$ for all $t \in \{0, ..., t'\}$, institution A is oversubscribed in each period. Therefore, Lemma 6 (ii) applies, and given $w^t = \bar{w}$ for $t \in \{1, ..., t'\}$, $v(C^{t'}) > 0$ in all equilibria, which proves the result. ■

Proof of Proposition 7:
For (i) note that if motivated workers value the ability-contingent reputation, then the model in this section reduces to the model presented in the general analysis – therefore the result follows from Corollary 3 and Proposition 5.

Take $\{p\}' \equiv \{p | \lambda_p > \bar{\lambda}\}$; note that $\{p\}'$ is non-empty since $\lambda^i \neq \lambda^j$ for some $i, j \in \{1, ..., p\}$. For (ii), a transition can be achieved with a wage path, $\{w_1^t, ..., w_p^t\}$, where:

$$w_p^t \begin{cases} = \overline{x}_p + v(C^0) & \text{if } p \in \{p\}' \\ = 0 & \text{if } p \notin \{p\}', \end{cases}$$

for $t = 1, 2$, and w_p^t is market-clearing given the myopic equilibrium for all p for all $t > 2$. Since only ability-types with average motivation above $\bar{\lambda}$ are employed in periods

1 and 2, but wages for these ability-types are high enough so that all workers apply, $C^2 = \sum^{p'} \lambda_p/|p'| > \bar{\lambda}$. That is, a transition to a reputation greater than $\bar{\lambda}$ is enabled by the fact that institution can achieve a high-motivation reputation by selectively raising wages for ability-types that have a relatively high proportion of motivated types. Given $v(\tilde{C}^2) > 0$, a transition to $\{w^*, C^*\}$ is achieved in all equilibria by the same argument as in the proof of Proposition 4 since Lemma 6 applies to each set $\{\tilde{C}_p\}$ individually. ■