Teamwork and Moral Hazard among Emergency Department Physicians*

David C. Chan†

March 23, 2014

Abstract

How does teamwork increase productivity? Considering teamwork as joint monitoring and management, I investigate this question by studying emergency physicians who work in two settings differing only in the extent that physicians manage work together: In a "nurse-managed" system patients are assigned by a triage nurse "manager," and in a "self-managed" system physicians decide among themselves which patients to treat. The self-managed system increases throughput productivity by 11-15%. Essentially all of this net effect can be accounted for by reducing a "foot-dragging" moral hazard, in which physicians prolong patient stays to appear busier and avoid getting new patients. Foot-dragging is sensitive to peer effects, suggesting that physicians in the same location have better information about each other. In the self-managed system, new patients are assigned more efficiently according to physician workload, suggesting a better use of information to assign patients.

JEL Codes: D20, L20, I12, M50

*I am very grateful to David Cutler, Joe Doyle, Bob Gibbons, and Jon Gruber for their guidance and support. David Bates and Josh Kosowsky provided invaluable institutional support. I also benefited from comments from Alberto Abadie, Jason Abaluck, Leila Agha, Josh Angrist, Amy Finkelstein, Nathaniel Hendren, Erin Johnson, Danielle Li, David Molitor, Michael Powell, Heidi Williams, and a large number of seminar audiences. I gratefully acknowledge support from the NBER Health and Aging Fellowship, under the National Institute of Aging Grant Number T32-AG000186; the Charles A. King Trust Postdoctoral Fellowship, the Medical Foundation; the MIT George and Obie Schultz Fund; and the Agency for Healthcare Research and Quality Ruth L. Kirschstein Individual Postdoctoral Fellowship F32-HS021044.

†Address: 117 Encina Commons, Room 215; Stanford, CA 94306. Phone: 650-725-9582. Fax: 650-723-1919. Email: david.c.chan@stanford.edu.

1 Introduction

Teams have become widespread in production across many industries. According to one summary, "teamwork has emerged in recent years as one of the most important ways in which work is being reorganized" (Delarue et al., 2008). Broad evidence suggests that teamwork and other human resource management technologies are associated with higher productivity (Ichniowski and Shaw, 2003; Bloom and Van Reenen, 2007; Hamilton et al., 2003). Despite this, economists have had little to say about how teamwork can increase worker productivity given the classic prediction that joint production leads to moral hazard (Alchian and Demsetz, 1972; Holmstrom, 1982).

In this paper, I consider teamwork as an organizational feature, distinct from joint production, that allows workers to monitor and manage each other.[1] While all workers in a firm arguably participate in joint production, they are not necessarily all engaged in teamwork with each other. Teamwork can *reduce* moral hazard by allowing workers to make use of better information about each other. This motivates my study of a natural experiment in which the same emergency department (ED) physicians work in two different organizational systems that differ only in the extent that physicians manage work together. In a "nurse-managed" system, two physicians in the same location ("pod") are individually assigned patients by a triage nurse "manager." In the second system, which I call "self-managed," the triage nurse first assigns patients to a pod shared by two physicians, who then decide between themselves who will care for each arriving patient. The self-managed treatment isolates the key ingredient of mutual management for teamwork.

While the effect of this treatment is intrinsically interesting, it is perhaps more important to understand the mechanisms behind such an effect. In the nurse-managed system, under asymmetric information between physician workers and the triage nurse manager, physicians may want to avoid being assigned more work by appearing busier than they are, keeping patients longer than necessary ("foot-dragging") in order to distort signals of their true workloads.[2] In the self-managed system, physicians may use better information about each other's true workloads to choose patients, thereby

[1] This concept has been highlighted in the psychology literature on management (Pallak and Perloff, 1986). Papers in the economics literature have shown that mutual monitoring may improve productivity (e.g., Knez and Simester, 2001; Mas and Moretti, 2009; Bandiera et al., 2010). This paper evaluates *mutual management* as a complementary mechanism for improving productivity in teamwork.

[2] In health care, equally colorful terms of "bed hiding" and "bed blocking" have been used to describe related phenomena that are prevalent and severe (Meisel and Pines, 2008; Proudlove et al., 2003), concerning enough for some observers to worry about the "collapse" of the U.K. National Health Service (Donnelly, 2013).

reducing foot-dragging relative to the nurse-managed system. However, physicians in the self-managed system may also seek to avoid work by waiting for their peer to pick patients first (a distinct moral hazard made possible by the self-managed system that I distinguish with the term "free-riding"). Finally, outcomes may differ through advantageous selection in the self-managed system, as physicians can choose patients according to either skill or availability.

Two sets of features of the empirical setting are advantageous for studying how teamwork improves productivity. First, the treatment, resulting from a simple rule change that received little attention, is limited in scope yet isolates a necessary component of teamwork: mutual management. One of two pods changed from a nurse-managed system to a self-managed system during the sample period, while the other pod always operated under a self-managed system. Moreover, the same health care providers work in both pods over time. Together, these features address worker selection and other unobservable characteristics at the organizational level that are non-trivial challenges in empirically studying teamwork and other management technologies.[3] Patient observations are frequent, with about one patient every nine minutes, while the time period spanning the change is six years, allowing me to confirm the conditional parallel trends assumption and use alternative methods of inference for the overall treatment effect, such as systematic placebo tests.

A second set of features unique to the ED identifies foot-dragging as a mechanism. The main feature uses patient flow to identify foot-dragging as a moral hazard in response to *expected future work*. Patient flow to the waiting room is highly unpredictable, even conditional on time categories used in scheduling. Further, although physicians can view patients in the waiting room via a computer interface, there remains substantial uncertainty in how these patients will be distributed to pods *ex post*. I thus identify foot-dragging as a response to variation in expected future work during patient care, separate from actual (current or future) work to a pod. In addition, I investigate the response to expected future work not only by pod and month but also by hour of the day, over which there is variation in the presence of a physician peer. Finally, based on the concept of foot-dragging as signal distortion, I test a prediction reminiscent of Milgrom and Roberts (1988): that signaled workload (i.e., the number of patients under a physician's care) should correlate more with patient assignment in the self-managed system.

[3]The approach of studying productivity in settings where the work environment is well understood is similar in spirit to other empirical studies in personnel economics (e.g., Ichniowski et al., 1997; Ichino and Maggi, 2000; Lazear, 2000; Hamilton et al., 2003; Bandiera et al., 2005; Mas and Moretti, 2009)

I find that physicians perform 11-15% faster in the self-managed system than in the nurse-managed system. The time a physician spends on a patient (i.e., the patient's length of stay) is particularly relevant because it defines a physician's contribution to the joint product of waiting times, a key determinant of patient satisfaction and health outcomes (e.g., Thompson et al., 1996). I find no other difference in individual quality or financial measures or in orders written, suggesting that physicians simply delay discharging their patients and provide no more or less care for them. I then more directly examine foot-dragging by using the assumption that, under this mechanism but not other mechanisms, lengths of stay should increase when physicians expect greater future work (but may not get it). Lengths of stay increase with expected future work in the nurse-managed system but not in the self-managed system. The effect is equal in magnitude to the overall difference between the two systems.

Further, I evaluate whether the presence of a peer reduces foot-dragging, as a joint test of better information and social incentives between peers in the same pod. I use the fact that the location of other physicians does not affect total work but could affect the ability of physicians to monitor each other's work. Thus, if physicians care about being seen engaging in moral hazard, the presence of a peer in the same pod could reduce foot-dragging. I find that the presence of a peer in the same pod substantially reduces foot-dragging in the nurse-managed system, in which physicians can monitor but not manage each other's work, relative to when there is another physician in the ED but not in the same pod.

Finally, I study patient assignment to test whether the self-managed system makes use of the better information between peers, through a prediction reminiscent of Milgrom and Roberts (1988): In the nurse-managed system, because physicians are tempted to foot-drag (i.e., distorting censuses – or the number of patients in their care – upward as signals of true workload), the triage nurse can be better off by discounting the informational value of censuses. In contrast, in the self-managed system, because physicians observe better information about each other's true workloads, censuses are less likely to be distorted and can be used more efficiently *ex post*. Consistent with this, I find that patient assignment is more negatively correlated with censuses in the self-managed system than in the nurse-managed system. I also study patient assignment in the pod switching to a self-managed system and find evidence of a transition period and enforcement against foot-dragging during this period, with higher-census physicians *more* likely to be assigned new patients.

Together, these findings suggest that mutual monitoring and management are key productivity-increasing components of teamwork. These results relate to two main strands of literature. First, a large literature has shown substantial variation in productivity across firms (Syverson, 2011). High-level managers (Bertrand and Schoar, 2003) and bundles of management practices (Ichniowski and Shaw, 2003; Ichniowski et al., 1997; Bloom and Van Reenen, 2007; Bloom et al., 2013) have been linked to productivity differences. However, features of management and organizational structure are difficult to describe and are rarely isolated. This study adds to the literature by isolating an important feature of teamwork and by illuminating mechanisms behind its effect.

Second, research has shown that monitoring can improve efficiency (Nagin et al., 2002; Duflo et al., 2013). Monitoring, however, is often considered external to the workers, even though workers likely have better information about their peers than managers or professional auditors. To this point, another literature on social incentives has shown that workers can behave more efficiently when they know their peers can observe them simply because they care about what their peers think (Mas and Moretti, 2009; Bandiera et al., 2009, 2005; Kandel and Lazear, 1992). Yet social incentives and mutual monitoring are often insufficient and sometimes even detrimental to productivity (Roy, 1952; Bandiera et al., 2005). This paper demonstrates joint management as an important ingredient for workers to use information between themselves for improving productivity.[4]

The remainder of the paper proceeds as follows. Section 2 outlines a simple model of asymmetric information in the assignment of work to explain how mutual monitoring and management between workers can reduce foot-dragging. Section 3 describes the ED institutional setting and data. Section 4 reports the overall effect of the self-managed system. Sections 5 and 6 discuss the main evidence for foot-dragging and its mitigation by organizational structure and the presence of peers. Section 7 explores patient assignment in the two systems over time. Section 8 concludes.

2 Theoretical Framework

In this section, I outline a simple model of asymmetric information between physicians and the triage nurse. The purpose of this model is to show how the self-managed system reduces

[4] A related third strand of literature deals with decentralizing decisions to workers who may be biased, rather than having workers communicate information to managers (Aghion and Tirole, 1997; Caroli and Van Reenen, 2001; Dessein, 2002; Acemoglu et al., 2007). However, this literature has dealt with decentralizing decisions to single workers, rather than teams of workers who can monitor each other.

foot-dragging and improves assignment efficiency relative to the nurse-managed system, formalizing the concept that teamwork improves productivity by "monitoring and managing work process and progress" (Pallak and Perloff, 1986).

I assume that in the nurse-managed system, the triage nurse cannot observe true physician workloads, although she would like to assign new work according to workloads.[5] Given that physicians prefer to avoid new work, they distort signals of true workload by prolonging patient lengths of stay (i.e., foot-dragging). At the same time, similar to Milgrom and Roberts (1988), I show that a triage nurse who takes this into account can be better off by committing to an *ex post* inefficient policy of ignoring signals, even though signals remain informative. In the self-managed system, however, if physician peers sometimes observe each other's true workloads, then they can also use that information to assign new work. This reduces the threat of foot-dragging and improves *ex post* assignment efficiency.

2.1 Stylized Pod Environment

Consider the following simple game of asymmetric information: Two physicians $j \in \{1, 2\}$ work in a single pod at the same time. They each have one patient, endowing them with low or high workload. In addition to the time that they take on their current patients, physicians also care about future work – a third patient – assigned to one of them. Physician utility is given by

$$u_j^P = -\left(t_j - \theta_j\right)^2 - K_P\left(\theta_j\right) \mathbb{I}\left\{\mathcal{J}\left(3\right) = j\right\}, \tag{2.1}$$

where t_j is the time that physician j keeps his initial patient, $\theta_j \in \left\{\underline{\theta}, \overline{\theta}\right\}$ is the workload entailed by his initial patient (where $\overline{\theta} > \underline{\theta} > 0$), $K_P\left(\theta_j\right) > 0$ is the cost of getting a potential third patient conditional on θ_j, and $\mathcal{J}\left(3\right)$ denotes the physician who gets the third patient.

Type $\underline{\theta}$ occurs with probability p. Types are never observed by the triage nurse, but with probability ψ, peers observe each other's types. In contrast, the number of patients of each physician (his census, $c_j \in \{0, 1\}$) is public information at any time. The action that each physician takes

[5] In practice, managers can know a lot about the status of their workers. In particular, the triage nurse observes patients before assigning them to physicians, and over time, she also knows the average productivity of each physician, as well as how this productivity might differ for different patient types. All that this model requires is that workers are more likely than the manager to observe each other's workloads.

is t_j. Absent any strategic behavior, each physician would like to discharge his patient at $t_j = \theta_j$, which I assume is socially optimal and generically captures all concerns of care (e.g., patient health and satisfaction, malpractice concerns, physician effort).[6]

The physician assigned the new patient incurs a cost, which depends on his initial workload, θ_j. I specify this cost as $K_P(\underline{\theta}) = \underline{K}_P$ and $K_P(\overline{\theta}) = \overline{K}_P$, where $\overline{K}_P > \underline{K}_P > 0$. This reflects the idea that neither physician would like to get the new patient, but that it is more costly for a physician with high workload, e.g., in terms of additional effort or worse outcomes for this new patient.[7]

The timing of the game is as follows:

1. At time $t = 0$ physicians each receive one patient, discovering $\theta_j \in \{\underline{\theta}, \overline{\theta}\}$.
2. Physicians simultaneously choose how long they will keep their patients, t_j.
3. With probability $\psi > 0$, physicians observe each other's θ_j.
4. Exactly one patient will arrive with uniform probability distributed across the time interval $t \in [\underline{\theta}, \overline{\theta}]$. Upon arrival, this new patient is assigned to a physician by the triage nurse (in the nurse-managed system) or the physicians themselves (in the self-managed system).
5. Physicians complete their work on the one or two patients under their care and end their shifts. They receive payoffs given in Equation (2.1).

This model highlights the tension between using signals (censuses c_j) of private information (types θ_j) for patient assignment and the fact that these signals can be distorted (through t_j). Physicians prefer to avoid new work (through $K_P(\theta) > 0$), but otherwise I assume that physicians have no incentive to keep patients longer than socially optimal. The triage nurse does not observe the physicians' types, but physicians observe each other's types with probability ψ.

[6] I thus abstract away from any principal-agent problem between physician and patient or between physician and ED in this term, other than the cost incurred by the new patient, discussed below.

[7] It is possible that $K_P(\theta) < 0$, e.g., when the professional gratification or financial incentive for seeing new patients is great. It is easy to show that if $0 > \overline{K}_P > \underline{K}_P$, then physicians having $\overline{\theta}$ will try to mimic having $\underline{\theta}$ (i.e., we should have the opposite of foot-dragging).

2.2 Nurse-managed System

In the nurse-managed system, the triage nurse assigns the new patient to a physician. In my baseline model, I assume that physicians cannot report their types or anything else to the triage nurse, but that the triage nurse can credibly commit to an assignment policy prior to physicians receiving their patients.[8]

The triage nurse's utility is

$$u^N = -D \sum_{j\in\{1,2\}} (t_j - \theta_j)^2 - K_N\left(\theta_{\mathcal{J}(3)}\right). \tag{2.2}$$

D is an indicator that allows the triage nurse to care about the treatment times of the first two patients as outcomes (if $D = 1$). Remember that $t_j = \theta_j$ is socially optimal and that this is universally agreed upon. I specify $K_N(\underline{\theta}) = 0$ and $K_N\left(\overline{\theta}\right) = \overline{K}_N$, where $\overline{K}_N > 0$. This represents that it is managerially preferable to assign new work to a worker with lower workload, e.g., because that worker is able to handle the new work in a more timely or higher quality manner. I do not restrict the the value of $\overline{K}_N$ relative to $\overline{K}_P - \underline{K}_P$.[9]

At $t = 0$, the triage nurse commits to an assignment policy function $\pi(c_1, c_2)$, with censuses $c_j \in \{0, 1\}$. To simplify the analysis, I impose a symmetric policy function with $\pi(0,0) = \pi(1,1) = \frac{1}{2}$ and $\pi \equiv \pi(0,1) = 1 - \pi(1,0)$. That is, when both physicians have equal censuses, the triage nurse should have no preference to send the new patient to one physician or the other, since she has no other information about who has lower workload at that time. Note that $\pi = 1$ represents what I mean by *ex post* efficiency, since the triage nurse can infer that if $c_j = 0$ and $c_{-j} = 1$, then j certainly had the lower workload.

I use a Perfect Bayesian Equilibrium as the equilibrium concept. In equilibrium, the triage nurse chooses the optimal assignment policy, summarized by $\pi^* \equiv \pi^*(0,1)$, given physician discharge strategies $\underline{t}^*$ and $\overline{t}^*$ for initial patients of type $\underline{\theta}$ and $\overline{\theta}$, respectively. Given π^*, physicians choose optimal discharge strategies $\underline{t}^*$ and $\overline{t}^*$.

[8] In Appendix A-2, I consider two alternative scenarios: (1) the pure signaling game which allows neither physicians to report their types nor the triage nurse to commit to an assignment policy, and (2) the mechanism design game (without transfers) which allows physicians to report types and the triage nurse to commit to a policy. Results are similar, but not surprisingly, reporting and commitment as additional capabilities both improve efficiency.

[9] Note also that if $D = 0$, then it does not matter what value $\overline{K}_N$ takes, as long as it is some positive number.

Proposition 1. *In the Perfect Bayesian Equilibrium for the nurse-managed system, physicians with $\underline{\theta}$ and $\overline{\theta}$ discharge their patients at $\underline{t}^* > \underline{\theta}$ and $\overline{t}^* = \overline{\theta}$, respectively, and the triage nurse assigns the new patient to the physician with census 0, when the other physician has census 1, with some probability π^* between $\frac{1}{2}$ and 1.*

First note that the triage nurse will never want to send the new patient with greater probability to a physician with $c_j > c_{-j}$. So physicians with $\overline{\theta}$ will never want to mimic those with $\underline{\theta}$ and will choose $\overline{t}^* = \overline{\theta}$. But physicians with $\underline{\theta}$ have some reason to mimic having $\overline{\theta}$. For a given $\pi \equiv \pi(0,1)$ previously chosen by the triage nurse, the optimization problem of physicians with $\underline{\theta}$, $\max_{\underline{t}_j} \mathbb{E}\left[u_j^P\left(\underline{t}_j; \pi, \underline{\theta}\right)\right]$, yields

$$\underline{t}^* = \underline{\theta} + \frac{\underline{K}_P}{2\left(\overline{\theta} - \underline{\theta}\right)}\left(\pi - \frac{1}{2}\right). \tag{2.3}$$

As long as the triage nurse is more likely to send the new patient to a physician she believes has lower workload (i.e., $\pi > \frac{1}{2}$), physicians with $\underline{\theta}$ will foot-drag to at least temporarily mimic those with $\overline{\theta}$.[10]

The triage nurse commits to π^* that maximizes her expected utility given $\underline{t}^*$ and $\overline{t}^*$. Substituting (2.3) into her expected utility and solving the first-order condition yields π^*. For simplicity, I present the expression for π^* if $D = 0$:

$$\pi^* = \frac{1}{2} + \frac{\left(\overline{\theta} - \underline{\theta}\right)^2}{\underline{K}_P}. \tag{2.4}$$

Equation (2.4) shows that the nurse's choice of π depends on the cost of getting the new patient for the *physician with lower workload*, because of his temptation to foot-drag. Even if she wants to optimize only the assignment of the third patient, committing to $\pi^* < 1$ may improve her expected utility, which is similar to Milgrom and Roberts' (1988) result that managers can be better off if they commit not to listen to subordinates who could undertake costly "influence activities." This commitment increases triage-nurse utility by decreasing foot-dragging.[11]

The assumption of a single patient arriving in the interval $t \in \left[\underline{\theta}, \overline{\theta}\right]$ is convenient for representing

[10] A nice feature of this simple two-type model is that the first-order condition does not depend on what the peer's type or strategy is, because there is only one patient each and thus one policy parameter π. Keeping this initial patient longer by dt decreases the likelihood of getting the new patient by $(\pi - 1/2)\, dt / \left(\overline{\theta} - \underline{\theta}\right)$ regardless of the peer's census. For this reason, parameters like p do not matter. This is shown in detail in Appendix A-2.1.

[11] I show this in Appendix A-2. In Appendix A-2.3, I also show that π^* is even lower if $D = 1$.

the temptation of moral hazard for physicians with $\underline{\theta}$. However, in practice there are of course many new patients, and I identify foot-dragging as the response of lengths of stay to the flow of expected future work, defined in terms of *numbers* of patients arriving at the ED triage. To capture this intuition, I can extend the model by changing the interval over which the single patient is expected to arrive, which results in replacing the interval $\overline{\theta} - \underline{\theta}$ in the denominator of Equation (2.3) with some $\Delta t \lesseqgtr \overline{\theta} - \underline{\theta}$, as long as $\underline{t}^*$ is an interior solution. I show details in Appendix A-2.5, but the intuition is straightforward. With an infinite flow of patients to the ED (as $\Delta t \to 0$), physicians should expect to get a new patient the minute they discharge one. With no expected future patients (as $\Delta t \to \infty$), there is no incentive to foot-drag.

2.3 Self-managed System

For the self-managed system, I assume the same physician utilities and information structure as in the nurse-managed system. The only difference is that the two physicians, not a triage nurse, are responsible for deciding who gets the new patient. Physicians choose both t_j and an action that determines the assignment of the new patient. I consider two microfoundations of this assignment action in the self-managed system, both continuing the baseline assumption that physicians cannot report their types.[12] In one microfoundation, each physician may only decide whether to choose an unattended patient at each point in time. Alternatively, physicians may commit to an assignment policy that uses censuses and, with probability ψ, observations of true workload. I present a brief discussion of results below; details and derivation are in Appendix A-3.

Proposition 2. *Consider the Perfect Bayesian Equilibrium for the self-managed system. If physicians cannot commit to a policy function, there will be no foot-dragging or* ex post *inefficient assignment, but some free-riding if $\psi < 1$. If physicians can commit to a policy function, and if $\psi > 0$, $\overline{K}_P - \underline{K}_P \geq \overline{K}_N$, and $D = 1$, then there will be less foot-dragging and more* ex post *efficient assignment, relative to the nurse-managed system, and no free-riding.*

[12] In Appendix A-3.3, I allow physicians to report their types to each other. This mirrors the case of physician reporting in the nurse-managed system, in Appendix A-2.4. In both cases, reporting improves efficiency.

Without physician commitment to an assignment policy, physicians play a war of attrition (e.g., Bliss and Nalebuff, 1984), both incurring a cost by the new patient remaining unattended, which reflects that ED patients generally have time-sensitive conditions, specified with an expansion of (2.1) in Appendix A-3.1. If types are observed and different, the physician with $\underline{\theta}$ will choose the new patient immediately upon arrival by subgame-perfect reasoning similar to Rubinstein (1982). On the other hand, if they do not observe each other's types, or if they observe that they have the same type, then they leave the new patient unattended with some probability at each point in time. In any case, physicians with $\underline{\theta}$ do not foot-drag, because it never reduces the chances of getting the new patient.[13] In the second microfoundation, physicians commit to an assignment policy that they follow if they cannot observe each other's true workload. With probability ψ, physicians can use each other's observed types to assign the new patient, which lowers the attractiveness of foot-dragging. They can afford to commit to an assignment policy with greater *ex post* efficiency than the triage nurse's in the nurse-managed system, primarily because foot-dragging is less of a threat with more information on true workloads.[14]

In summary, under both microfoundations (with reasonable parameters in the commitment case), physicians foot-drag less, and the new patient is given more often to a physician with lower workload.[15] I consider both microfoundations because the truth likely lies in between: Physicians are not forced to see patients in the self-managed system, but they very likely have a strong cultural norm that quickly assigns patients in order to prevent patients from waiting in the pod unattended for too long. The model also predicts little free-riding with sufficient physician information (high ψ), commitment, or costs of leaving patients unattended.

The key point is that the threat of moral hazard is reduced by mutual monitoring and mutual management in the self-managed system. It is useful to contrast this with social incentives, which also reduce moral hazard with mutual monitoring, because peers incur a social cost when seen

[13] In Appendix A-3.1.4, I discuss that there is some foot-dragging with continuous types, but that foot-dragging will still be less than in the nurse-managed system with no triage nurse commitment, in Appendix A-2.3.

[14] In Appendix A-3.2, I show that another reason for improved *ex post* assignment inefficiency is that physicians are likely to care relatively more about inefficient assignment than the triage nurse (i.e., $\Delta K_P > \overline{K}_N$), since the cost of inefficient assignment is scaled relative to treating their own patients, while the triage nurse scales this relative to treating all patients.

[15] It is possible that physicians may collude and act as a single physician (Roy, 1952). This is not explicitly considered in this model, since there are only two physicians, of whom one *must* be assigned the new patient. This could be allowed in a model with a third physician in a different location who could be colluded against. However, even with full collusion in this augmented model, foot-dragging would be equal in the two systems; with less than full collusion, foot-dragging would be lower in the self-managed system.

engaging in moral hazard (Kandel and Lazear, 1992).[16] Note that self-management can improve efficiency without social incentives and that social incentives do not require mutual management. Nevertheless, because self-management and social incentives both depend on information between peers, I examine peer effects on foot-dragging in Section 6, as a useful test of both social costs and information between peers ($\psi > 0$).

3 Institutional Setting and Data

I study a large, academic ED with a high frequency of patient visits, greater than 60,000 visits per year (or 165 visits per day), with a total of 380,699 visits over six years. The ED is an especially appropriate setting to study the joint production of throughput. Because the time spent waiting for ED care is believed to adversely affect both patient satisfaction and health (Thompson et al., 1996), improving throughput is a top priority for many EDs and the focus of ED management consulting (McHugh et al., 2011).

My primary outcome, length of stay, measures each physician's individual contribution to jointly produced waiting times. While waiting times are affected by a host of factors beyond the control of a single physician, length of stay – defined as the time between patient arrival at the pod and the physician's discharge order – captures the component most directly controlled by the physician and is unaffected by inpatient bed availability.

Physicians have substantial discretion in a specific patient's length of stay, because of asymmetric information between physicians and managers and because other dimensions of productivity are important. After assuming the care of a patient, a physician may encounter clinical situations that warrant a longer length of stay to ensure quality care. To measure the quality of care, I focus on three prominent outcomes (e.g., Schuur and Venkatesh, 2012; Forster et al., 2003; Lerman and Kobernick, 1987). Thirty-day mortality is perhaps the most unambiguous but occurs in only 2% of the sample. Hospital admission is a resource-intensive discharge option, which may substitute for appropriate care in the ED. Bounce-backs, defined as return visits from home within 14 days,

[16] It is straightforward to modify physician utility in (2.1) such that their expected utility includes a term $\psi S(\cdot)$, where $S(\cdot)$ are social costs incurred conditional on being seen foot-dragging, that reduces foot-dragging.

represent the complementary quality concern of premature home discharge.

I also consider patient-level revenue and costs that accrue to the ED and hospital. For revenue, I use Relative Value Units (RVUs), which are units of physician billing for services that scale directly to dollars and reflect the intensity of care provided to a patient.[17] For costs, I use total direct costs for each patient encounter, including any costs incurred from a resulting hospital admission. Finally, I use detailed data on physician orders – approximately 13 per visit, including orders for nursing, medication, laboratory, and radiology – to shed light on the process of patient care. I do not observe the time that a physician officially signs up for a patient, but as a proxy for this, I use the time that the physician writes his first order.

3.1 Organizational Systems

A small feature in the process of patient assignment distinguishes two organizational systems, which are the focus of this paper (Figure 1). After patients arrive at the waiting room, or "triage," a triage nurse decides where and when to send them. In a "nurse-managed" system, beds within an ED location ("pod") are owned by one of potentially two physicians. The triage nurse therefore serves as a manager by directly allocating new patients to physicians.[18] In another "self-managed" system, two physicians in the same pod share the beds and are jointly responsible for dividing work sent to the pod by the triage nurse.

The assignment of patients to other health care providers, i.e., nurses and residents, does not differ between the two systems; in both systems, nurses are assigned patients, and residents choose patients. Regardless of the system of assignment, a single ED physician is responsible for the care of each patient once assigned, and ED physicians rarely confer with colleagues on patient care.[19] Financial incentives are also held constant: Physicians are paid a salary plus a 10% productivity bonus based on clinical productivity (measured by Relative Value Units, or RVUs, per hour).

[17]The current "conversion factor" is $34 per RVU, and the average ED patient is billed for 2.7 RVUs of ED care, resulting in about $6 million in yearly revenue for this particular ED.

[18]The assignment of patients by nurses or non-medical staff is the predominant system of work assignment in hospital and ED settings. This of course is "management" in a very limited sense, as the triage nurse cannot hire or fire physicians nor set financial incentives. In some settings, these non-medical "managers" may have no discretion but merely follow rules.

[19]Unlike the intervention in Hamilton et al. (2003), the self-managed system involved no team incentives or expectations that physicians collaborate in patient care. Rather than a benefit to physicians in helping each other work faster, doing so in the self-managed system may induce the triage nurse to assign more work to the team. More significant collaboration and other human resource management practices at least in some settings could further improve productivity (e.g., Hamilton et al., 2003; Ichniowski et al., 1997).

Information about patients cared for by each physician is available to all physicians in the ED and the triage nurse from a computer interface (Figures B-7.1 and B-7.2). A salient summary statistic of workload is a physician's census, or the number of patients being cared for by him, but this of course can be distorted by foot-dragging. Additional information available to physicians working in the same pod but likely not the triage nurse (or physicians outside of the pod) stems from closer observation of peer behavior (e.g., Does the peer appear busy? Is he talking to nurses?) and of events relevant to patient status (e.g., Are staff frequently in the room of a peer's patient?).

3.2 Features for Identifying the Overall Effect

The ED has two pods: "Alpha" and "Bravo." Alpha pod has always had a self-managed system. In contrast, in March 2010, Bravo pod switched from a nurse-managed system to a self-managed one. The regime change in Bravo pod resulted from a simple intervention in which beds that physicians previously owned became shared. Prior to the change, Bravo bed ownership was assigned by non-overlapping regions, Bravo-1 and Bravo-2. The change did not affect the physical layout of beds but formally allowed physicians to see patients in any Bravo bed.

The reason for this switch was to allow greater flexibility in patient assignment within pod, as Bravo became increasingly busy over time. According to interviews with ED administrators and physicians, the switch was not considered a significant change in organizational structure, and overall implications for efficiency were not apparent.[20] Schedules and staffing for providers and algorithms for patient assignment to beds, nurses, and residents remained unchanged. In fact, preassigned shifts retained vestigial Bravo-1 and Bravo-2 labels until the next academic year in July 2010. Although the switch was announced in January 2010, there was no formal pilot period. Importantly, there were no concurrent changes, including other interventions to promote teamwork.[21]

As an important time-invariant difference, Alpha pod has always been opened 24 hours, while Bravo pod has always closed at night. As a result, patients who need to stay longer, either because they are sicker or have conditions that might make discharge difficult (e.g., psychiatric patients),

[20]In fact, in May 2011, the ED moved both pods to the nurse-managed system, only to discover later that it significantly reduced efficiency. They reversed this organizational change in January 2012. I do not use data after May 2011 to study organizational systems at the pod level, because both pods switched systems and because May 2011 also witnessed the opening of a third pod, Charlie.

[21]In contrast, for example, the May 2011 switch was trumpeted as a major reorganization and was accompanied by a pilot period and team-building workshops.

have tended to be sent to Alpha pod. Closing Bravo every night may also prompt earlier discharges for patients in the pod as it nears closing. In addition to this difference, as the ED volume increased over time, Bravo – the traditionally less-intensive pod – received more patients and more time-intensive patients over time.[22]

Providers work in both pods over time, and the vast majority of providers are therefore observed in both organizational systems. I observe 92 physicians, 364 nurses, and 986 residents in the data. Among these, 75 physicians, 334 nurses, and 882 residents, comprising 11,865 unique physician-nurse-resident trios, are observed in both organizational systems.[23]

For each visit, I observe patient demographics, including age, sex, race, language, zip code of residence, and insurance status. I capture patient severity by the Emergency Severity Index (ESI), a five-level ED triage algorithm based on a patient's pain level, mental status, vital signs, and medical condition (Tanabe et al., 2004); in some specifications, I also use the time spent in triage relative to other patients as a measure of the triage nurse's judgment of the patient's (lack of) urgency. Diagnoses are recorded as International Statistical Classification of Diseases, Version 9, (ICD-9) codes and summarized for analysis as Elixhauser indices, which are 30 dummies for relevant comorbidities, such as heart failure or diabetes (Elixhauser et al., 1998).

3.3 Features for Identifying Foot-dragging

ED physicians can form real-time expectations of future work based on patients in the waiting room who have not yet been assigned a bed. Two features of the setting allow me to identify foot-dragging by examining physician responses to shocks of expected future work. First, patient flow in the ED is highly unpredictable, even conditioning on rough time categories used for physician scheduling (Appendix B-4.1). Physician schedules are determined one year in advance, and physicians are only able to request rare specific shifts off, such as holidays or vacation days.[24] General preferences, such as whether they would like to work at night, may be voluntarily stated but not

[22]For example, patients in Alpha were older and had more severe conditions, but this difference diminished over time. I show differences in observable patient characteristics between the pods over time in Appendix B-1.1. As I will discuss later, this works against finding improvements in productivity in Bravo.

[23]Essentially all providers who do not work in both systems either are occasional moonlighters or represent errors in recording the correct provider. For example, the number of visits corresponding to median resident is 1,525, while this number is 17 for residents who are observed to work in only one system.

[24]Shift trades are also exceedingly rare, about less than one per month, or <1% of the number of shifts. Results are robust to dropping traded shifts. Per ED administration, shifts are not assigned with peers in mind.

honored fully, and all physicians are expected to be open for shifts at all times of the day and days of the week. Once working on a shift, physicians cannot control the volume and types of patients arriving in the ED.[25]

Conditional on the month-year, day of the week, and hour of the day, I find that physicians are exposed to similar patients types arriving at their pod and patient numbers arriving at the ED. Table 1 shows this descriptively for physicians with above- and below-median productivity, defined by average lengths of stay, and more formal results are shown in Appendix B-5. This suggests that I can isolate shocks in expected future work that physicians cannot select or control.[26]

Second, patients in the waiting room are yet to be assigned to pods. The distribution of patients between pods can further vary widely, within a given volume of patients arriving at the ED. Depending on the time interval, the correlation between overall volume and pod-specific volume is at most 0.21, reflecting discretion by the triage nurse and stochastic discharge times creating differences in bed availability at the pod level (Appendix B-4.2). This variation allows me to separate expectations of future work (patients arriving at the ED but not yet assigned to a pod) from actual future or current work (patients who have or will be assigned to the pod a physician is working in). This feature is important for separating foot-dragging from other mechanisms in which physicians use teamwork to improve how they handle patients assigned to their pod.

As additional support for the hypothesis that self-management improves productivity by reducing foot-dragging, I use two other features of the data. First, although two physicians generally occupy a pod, depending on schedules, a physician may be alone in the pod or in the ED during certain times of the day. This allows me to test whether physicians differ in their response to expected future work, in the same pod and in the same shift, but when there may not be a peer present to monitor them. Second, I use observations on the assignment of new patients, conditional on physician censuses, as a test of the relative quality of censuses as signals of workload under the two organizational systems.

[25] Physicians may rarely (<1-2% of operating times) put the ED on "divert" for up to an hour when the flow of patients is unusually high and the entire ED lacks capacity to see more patients. Even when this happens, this only affects some ambulances (which as a whole constitute 15% of visits) carrying serious emergencies, as opposed to the majority of patients, some of whom walk in. ED flow is largely unaffected.

[26] This feature is actually unnecessary if I assume physicians have fixed differences in productivity that I can control for, because I observe physician identities. However, it addresses more complicated threats to identification, such as if physicians know in advance that they may be less productive on some days and would prefer to be in a certain organizational system with a certain expected level of work.

4 Overall Effect of the Self-managed System

In this section, I estimate the overall effect of the self-managed system on a given team of providers for a given patient. I specifically ask the following: If the same patient and providers were assigned to each other in a different organizational system, what would their outcomes be? I can control for pod-specific time-invariant unobservable differences by the fact that I observe one of the two pods (Bravo) switching from a nurse-managed system to a self-managed one. I can also control for providers because I observe the vast majority of providers – physicians, residents, and nurses – working in both pods over time.

As my baseline analysis, I estimate the following equation:

$$Y_{ijkpt} = \alpha Self_{pt} + \beta \mathbf{X}_{it} + \eta \mathbf{T}_t + \zeta_p + \nu_{jk} + \varepsilon_{ijkpt}, \tag{4.1}$$

where outcome Y_{ijkpt} is indexed for patient i, physician j, resident-nurse k, pod p, and arrival time t. The variable of interest in Equation (4.1) is $Self_{pt}$, which indicates whether pod p had a self-managed system at time t. It also controls for patient characteristics $\mathbf{X}_{it}$, time categories $\mathbf{T}_t$ (for month-year, day of the week, and hour of the day), pod identities ζ_p, and physician-resident-nurse trio identities ν_{jk}.

Interpreting α in Equation (4.1) as the causal effect of the self-managed system relies on the familiar assumption of parallel trends, conditional on rich observed characteristics described in Section 3.2, for patients sent to Alpha versus Bravo over time.[27] In Table 2, I estimate several versions of (4.1), including progressively more controls for patient characteristics. The estimate for the effect of self-managed teams on log length of stay remains stable (and slightly *increases* in magnitude) from -11% to -13% upon adding a progressively rich set of controls. This is consistent with the fact that more time-intensive patients were sent to Bravo pod over time (Appendix B-1). Adding controls for and interactions with peer characteristics does not change results (Appendix B-2). Adding pod-specific time trends (i.e., adding $\gamma_p t$) to Equation (4.1) yields a slightly larger effect of -15%. Note finally that, since work can be sent to either pod, an improvement in throughput

[27] As discussed in Section 3, there were secular trends between the two pods. Specifically, more and sicker patients were sent to Bravo pod over time, and new nurses generally spent more time in Bravo pod to fill the need of higher volume. This creates a positive bias, which is opposite in sign of the estimated effect on length of stay. I show trends in observable characteristics and unconditional results in Appendix B-1.

in one pod would likely shorten lengths of stay in the other pod by lessening congestion. This violation of the Stable Unit Treatment Value Assumption (SUTVA) also biases the estimated size of the causal effect downwards (Rubin, 1980; Crepon et al., 2013).

This overall effect represents a significant difference in length of stay due to a simple organizational change, in which physicians assign work among themselves, while the physicians themselves and financial incentives were held fixed. As a comparison, this effect is roughly equivalent to one standard deviation in physician productivity fixed effects: Physicians who are one standard deviation faster than average have lengths of stay that are about 11% shorter. Given average lengths of stay, the self-managed-system effect is equivalent to a reduction in lengths of stay by 20-25 minutes per patient, and under simple assumptions, it represents a $570,000 yearly savings to this single ED.[28] For this particular ED, the cost of implementing the organizational change associated with this effect was essentially free.

While I find a significant effect of self-managed teams on length of stay, I find no statistically significant effect on quality outcomes (30-day mortality, hospital admissions, 14-day bounce-backs) or financial/utilization outcomes (RVUs, total direct costs) at the 5% level, shown in Appendix Table B-8.1. Alternative mechanisms of free-riding and advantageous selection could affect the quality of care and utilization, because they mean that specific patients either are being made to wait for care or are seen by physicians who are better suited (or more available) to see them. In contrast, under pure foot-dragging, only the *discharge* of patients is delayed in order to prevent more work. Foot-dragging should not result in different quality or utilization between the self-managed and nurse-managed systems, because any impact through waiting times would be shared by *all* patients in the ED. Thus, the lack of effect on quality, revenue, and utilization between the two systems is more consistent with foot-dragging than other mechanisms such as better matching.

Any statement on quality, however, is limited by relatively imprecise estimates for mortality and bounce-backs. While I can rule out a 0.8% increase in mortality, this is relatively large compared to the baseline 2.0% mortality. More direct evidence of foot-dragging is shown in the next section. Nevertheless, estimates for admissions, RVUs, and costs are quite precisely estimated. For example,

[28] For this back-of-the-envelope calculation, I simply assume that ED patient volume is exogenous and that the ED is able to reduce the number of physician-hours, given improved throughput, to meet the volume. By allowing more patients to be seen for a given number of physician-hours, the $570,000 yearly saving to this ED derives from $4.4 million yearly spending in physician hourly salaries (26,280 physician-hours per year at about $167 per physician-hour). This gain ignores reduced waiting times and improved outcomes shared by all ED patients.

given the current dollar conversion of about \$34 per RVU, the average ED patient represents about \$92 in revenue, while the effect of self-management on revenue is only -\$0.51 (95% CI -\$2.38 to \$1.36). With no change in revenue or costs per patient, delaying the discharge of patients thus unambiguously decreases productivity from a financial perspective.

Table B-8.1 also reports the effect on the time to the first physician order, which is an upper bound for the time to being chosen by a physician. Significant free-riding would imply a positive coefficient for the self-managed system with respect to this proxy. However, the effect of the self-managed system on this measure is not significantly different from 0 and slightly negative.

The effect on length of stay due to Bravo's regime change to a self-managed system can also be seen graphically. Figure 2 shows month-year-pod fixed effects over time for the two pods estimated by this equation:

$$Y_{ijkpt} = \sum_{m=1}^{M} \sum_{y=1}^{Y} \alpha_{myp} I_{t \in m} I_{t \in y} + \beta \mathbf{X}_{it} + \tilde{\eta} \tilde{\mathbf{T}}_t + \nu_{jk} + \varepsilon_{ijkpt}, \tag{4.2}$$

where the parameters of interest α_{myp} are fixed effects for each month, year, and pod interaction; $I_{t \in m}$ and $I_{t \in y}$ are indicator functions for t belonging to month m and year y, respectively; and $\tilde{\mathbf{T}}_t$ is a revised vector of time categories that only includes day of the week and hour of the day. In Figure 2, there is a persistent decrease for Bravo around the regime change that is consistent with my baseline estimates in Table 2 that self-managed teams in Bravo decreased patient lengths of stay. In addition, Figure 2 shows that trends in log length of stay in the two pods are roughly conditionally parallel, which is the identifying assumption for Equation (4.1).

An issue that arises in difference-in-differences estimation is the construction of appropriate standard errors for inference (Bertrand et al., 2004).[29] My baseline specification clusters standard errors by physician, suggesting an experiment sampling at the level of physicians, who are given shifts that translate to pods and organizational systems. This thought experiment is supported by the fact that the same physicians work in both pods before and after the regime change, and by evidence that physicians are exposed to quasi-random patients and peers, conditional on rough time categories (Appendix B-5).

[29] This issue is largely only relevant to standard errors for the overall effect. Specific mechanisms use additional variation. In particular, foot-dragging relies on the effect of due to exogenous variation in expected future work.

Still, there is the additional statistical issue of unobserved and potentially correlated pod-level shocks over time. Therefore, I consider two alternative thought experiments for inference, both of which exploit the long time series and can be understood by Figure 2. First, I address sampling variation at the pod level across time but with a more parametric form, assuming a month-year-pod shock that is correlated by a first-order autoregressive process across months within pod. Second, in the spirit of systematic placebo tests (Abadie et al., 2011, 2010) and randomization inference (Rosenbaum, 2002), I consider the thought experiment that, under the sharp null of no effect of self-managed teams, there should be no significant difference between my estimates and those obtained for placebo regime changes in both pods and other months. Detailed in Appendix B-3, both approaches yield a high degree of statistical significance.

5 Main Evidence of Foot-dragging

This section identifies the mechanism of foot-dragging with the following intuition: The expected gains to physicians by foot-dragging depend on expectations of future work. If no further patients arrive at the ED, then foot-dragging is not needed to prevent new work. But if there is an endless supply of patients waiting to be seen, then discharging a patient directly leads to having to see another one, and the incentive to foot-drag is extremely strong. I thus identify and quantify foot-dragging by increases in lengths of stay as expected future work increases.

The identifying assumption is that only the moral hazard of foot-dragging should increase lengths of stay with expected future work, holding actual work constant.[30] I therefore interpret increasing length of stay with expected future work as evidence of foot-dragging. I hold this interpretation regardless of organizational system, but I am also interested in comparing foot-dragging between organizational systems. Also, as outlined in Section 2 above, it is important to note that the costs of foot-dragging derive from both direct moral hazard and inefficient assignment that results from this moral hazard. Both of these are directly related to expectations of future work, and both are jointly determined in equilibrium. I do not separately identify these effects here, but I separately

[30]When both pods are open, patients who arrive at the ED while the index patient has just arrived at the pod may be sent to either pod. This allows me to separate expectations of future work (the number of patients arriving at the ED) from actual current or future work (the number of patients who arrive at the pod). I discuss this further below. In Appendix B-4.1, I show significant variation in patient volume even within time categories; in Appendix B-4.2, I show significant variation in pod-specific volume conditional on ED volume.

address assignment in Section 7.

For my baseline estimation of foot-dragging, I use equations of the following form:

$$\begin{aligned} Y_{ijkpt} &= \alpha_1 EDWork_t + \alpha_2 Self_{pt} \cdot EDWork_t + \alpha_3 Self_{pt} + \\ &\quad \beta \mathbf{X}_{it} + \eta \mathbf{T}_t + \zeta_p + \nu_{jk} + \varepsilon_{ijkpt}, \end{aligned} \tag{5.1}$$

for log length of stay Y_{ijkpt} for patient i, physician j, resident and nurse k, pod p, and time t. As before, $Self_{pt}$ indicates whether pod p at time t was self-managed, $\mathbf{X}_{it}$ controls for patient characteristics, $\mathbf{T}_t$ is a vector of time categories (month-year interaction, day of the week, hour of the day), ζ_p controls for the pod, and ν_{jk} controls for the provider-trio.

$EDWork_t$ represents expected future work, defined in two ways. First, I consider ED arrival volume, defined as the number of patients arriving at triage in the hour prior to patient i's arrival at the pod. The arrival of these patients is not controlled by physicians. They are seen by physicians via the computer interface, but their ultimate destination is unknown. Second, I consider waiting room census, defined as the number of patients (the census) in the waiting room at the time of patient i's arrival at the pod. Although physicians presumably can affect the number of patients waiting, this is a more salient measure of expected future work since physicians can readily click on the computer interface to see this census. Since I control for time categories, I identify the response to idiosyncratic shocks of future work that are not expected at the time of scheduling but expected only in real time as patients arrive at the waiting room. The coefficients of interest in (5.1) are α_1, α_2, and α_3. A positive α_1 indicates that physicians increase lengths of stay as expected future work increases (i.e., they foot-drag) in the nurse-managed system, while a negative α_2 indicates that the self-managed system mitigates foot-dragging. Coefficient α_3 represents the effect of the self-managed system after controlling for foot-dragging.

Table 3 reports estimates for (5.1), using both measures of expected future work – ED arrival volume and waiting room census. With each additional patient arriving hourly at triage or waiting in triage, lengths of stay increase by 0.6% in the nurse-managed system. The estimate of foot-dragging in the nurse-managed system is equivalent to a length-of-stay elasticity of 0.10 with respect to expected future work.[31] The coefficient for the interaction between expected future work and

[31] I estimate this by using log measures of expected future work as $EDWork_t$. My preferred specification, shown in

the self-managed system suggests that this effect is entirely mitigated in the self-managed system. That is, an additional patient in either measure of expected future work does not affect lengths of stay in the self-managed system. After controlling for foot-dragging, the coefficient representing the effect of the self-managed system is statistically insignificant in all specifications and ranges from -1% to 3%. In addition to the baseline specification in (5.1), I also include a number of controls for physician workload with pod-level volume at the time of patient arrival. Results are robust to including these controls.

These results suggest substantial foot-dragging in the nurse-managed system and none in the self-managed system. Estimates are robust under both measures of expected future work – ED arrival volume and waiting room census. Controlling for actual current or future pod-level work does not change results, suggesting that increased lengths of stay are due to *expectations* of future work. Again, this distinguishes foot-dragging from other mechanisms of free-riding or advantageous selection. Mechanisms of free-riding and advantageous selection only apply to patients who can be chosen, who are in the same pod.[32] Finally, the remaining effect of the self-managed system is insignificant, suggesting that foot-dragging is quantitatively equivalent to the overall difference in performance between self-managed and nurse-managed systems.

Figure 3 plots log-length-of-stay coefficients for each decile of expected future work interacted with organizational system, estimated by

$$\begin{aligned} Y_{ijkpt} \;=\; & \sum_{d=2}^{10} \alpha_0^d \left(1 - \mathit{Self}_{pt}\right) D_d\left(\mathit{EDWork}_t\right) + \sum_{d=1}^{10} \alpha_1^d \mathit{Self}_{pt} D_d\left(\mathit{EDWork}_t\right) + \\ & \beta \mathbf{X}_{it} + \eta \mathbf{T}_t + \zeta_p + \nu_{jk} + \varepsilon_{ijkpt}, \end{aligned} \tag{5.2}$$

where $D_d\left(\mathit{EDWork}_t\right)$ equals 1 if EDWork_t is the d^{th} decile, measured either as ED arrival volume or waiting room census. The coefficients $\left\{\alpha_0^2, \ldots, \alpha_0^{10}; \alpha_1^1, \ldots, \alpha_1^{10}\right\}$ can be interpreted as the relative expected length of stay for patients in different organizational systems and under different states of

Table 3, does not take logs of expected future work because it is roughly normally distributed. However, results are qualitatively the same in this specification.

[32] Two potential exceptions of spillovers between pods are waiting for a radiology test or a hospital bed. However, I find no difference in foot-dragging between patients based on likelihood, estimated by patient characteristics, to receive radiology tests. Table B-8.2 also shows that radiology testing is unaffected by expected future work. The time spent waiting for a hospital bed is excluded from my measure of length of stay, since I record the time of the discharge *order*. In Table B-8.2, outcomes like admission, which are not supposed to be affected by foot-dragging, suggest spillovers from hospital congestion applying in *both* organizational systems.

expected future work, where the expected length of stay for patients in the nurse-managed system and under the first decile of expected future work is normalized to 0. As shown in Figure 3, lengths of stay progressively increase in the nurse-managed system as expected future work increases, which is consistent with the intuition that the incentive to foot-drag continues to grow as expected future work increases. The self-managed team has roughly the same expected length of stay as the nurse-managed team at low patient volumes, but its expected length of stay does not change with patient volume.

My measures of expected future work are likely to be noisy representations of physicians' true expectations of future work.[33] Therefore this estimate is a lower bound on true foot-dragging: It is biased downward to the extent that I do not capture true expectations of future work. In addition, I interpret any increase in length of stay with expected future work as foot-dragging. But it may reasonable to think that physicians in the absence of moral hazard should actually work *faster*, for example if they care about patients waiting too long in the waiting room. This is another sense in which my interpretation is a conservative benchmark: It assumes that, under no foot-dragging, there is either no attention to future work or no reason to work faster when future work increases. Note that since length of stay does not increase with expected future work in the self-managed system, foot-dragging relative to 0 and foot-dragging relative to self-managed teams are roughly the same in magnitude.

I also estimate the baseline equation for foot-dragging, Equation (5.1), for other outcomes and process measures: 30-day mortality, admissions, 14-day bounce-backs, RVUs, total direct costs, and a host of detailed process measures including laboratory, medication, and radiology orders. Shown in Appendix Table B-8.2, I find no differential effect of expected future work between the two systems on any of these outcomes or process measures. Some outcomes do reflect a slight effect of ED arrival volume through hospital congestion for *both* systems, such as on hospital admissions and total costs, which include costs incurred in admissions. Foot-dragging effects on process measures are tightly estimated and show that the care provided while foot-dragging is not substantively different between the two systems. This is also consistent with pure foot-dragging, which delays the time of patient discharge but does not increase the quality or content of medical care.

[33]Note also that I do not use expectations based on the usual volume for the time of the day or day of the week, which are absorbed by time fixed effects.

6 Peer Effects on Foot-dragging

Foot-dragging could be reduced by the presence of peers if physicians have more information about each other than the triage nurse has and if they also care about being seen foot-dragging. While physicians usually work in pods with a peer, during certain times on shift, physicians find themselves working without a peer. This allows me to explore peer effects on foot-dragging, as a joint test of more information between peers and social incentives.[34]

In the nurse-managed system, only the joint existence of social incentives *and* more information between peers can explain the dependence of foot-dragging on the location of coworkers. In addition, two similar analyses can serve as falsification tests for the identification of foot-dragging by increases in length of stay with expected future work. First, when a physician in a self-managed pod is without a peer, he is effectively in a nurse-managed system: Every patient who arrives is in fact assigned to him by the triage nurse. The physician should then exhibit foot-dragging behavior as if working in the nurse-managed system (and without a peer). Second, when there is only one physician in the entire ED, there is essentially no assignment problem. That physician is responsible for all patients who arrive at the ED. With no coworker to foot-drag against, physicians have no incentive for foot-dragging.[35]

In Table 4, I present results for regressions of the form

$$\begin{aligned} Y_{ijkpt} \quad &= \quad \alpha_1 EDWork_t + \alpha_2 NoPeer_{jt} \cdot EDWork_t + \alpha_3 NoPeer_{jt} + \\ &\quad\quad \beta \mathbf{X}_{it} + \eta \mathbf{T}_t + \zeta_p + \nu_{jk} + \varepsilon_{ijkpt}, \end{aligned} \tag{6.1}$$

for nurse-managed-team and self-managed-team samples separately. $EDWork_t$ is ED arrival volume, or the number of patients arriving at ED triage in the hour prior to the index patient's arrival at the pod, and $NoPeer_{jt}$ is a dummy for whether physician j has no peer in the same pod. I estimate

[34] Peer effects *on foot-dragging* are the effect of a peer interacted with expected future work; it isolates the effect of peers on foot-dragging moral hazard. In contrast, generic peer effects are simply the direct effect of a peer and could act through a variety of mechanisms, such as productivity spillovers. I discuss more generic peer effects in Appendix B-6, where I show results similar to Mas and Moretti (2009) (i.e., productive peers increase the productivity of physicians).

[35] Physicians can still foot-drag against future physicians, but this theoretically is no different at any other time. I cannot control for actual future work in this scenario, because all work that comes to the ED eventually goes to the same pod and physician. However, this would only bias measured foot-dragging upwards if the omitted volume of actual work is positively correlated with expected future work and increases length of stay.

Equation (6.1) only when there are at least two physicians in the ED, so that foot-dragging always entails a negative externality against a current coworker. Note that since I am interested in the interaction between peer presence and measures of patient flow, I can control for time categories, $\mathbf{T}_t$, of hour of the day, day of the week, and month-year interactions.[36]

I also perform the pooled regression

$$\begin{aligned} Y_{ijkpt} &= \sum_{s=1}^{4} 1\left(PeerState_{jt} = s\right)\left(\alpha_s EDWork_t + \delta_s\right) + \\ & \quad \beta \mathbf{X}_{it} + \eta \mathbf{T}_t + \zeta_p + \nu_{jk} + \varepsilon_{ijkpt}, \end{aligned} \tag{6.2}$$

which estimates the degree of foot-dragging with coefficient α_s for each of four peer states $s \in \{1, \ldots 4\}$: alone in a pod but not alone in the ED, with a peer in the nurse-managed system, with a peer in the self-managed system, and alone in the ED.

Results in Table 4 are consistent with previously estimated coefficients for the increase in length of stay with expected future work, shown in Table 3, in both systems when a peer is present. When a peer is not present, however, length of stay increases much more quickly with expected future work. Estimates in the first two columns suggest that, without a peer present, the response to expected future work quintuples in the nurse-managed pod and increases in the self-managed pod (but effective nurse-managed system) to almost triple the magnitude as in the nurse-managed system with a peer. Results from the pooled regression in Equation (6.2), shown in the third column of Table 4, confirm this and show that physicians do not increase lengths of stay with expected future work when they are alone in the ED.

These results suggest that physicians reduce foot-dragging when a peer is present, which is consistent with social incentives and the fact that peers observe true workload better than do physicians in a different location.[37] Additionally, results from the two falsification tests, in the second and third columns of Table 4, support the interpretation of increases in length of stay with expected future work as foot-dragging moral hazard. First, physicians working without a peer in an officially self-managed pod (but effectively nurse-managed system) engage in foot-dragging to triple

[36] In addition, due to scheduling changes, there exists variation in the times in which a physician is alone in a pod.

[37] In Appendix B-6.2, I show that peer effects on foot-dragging can also depend on the relationship with the peer present. In particular, working with a senior peer diminishes foot-dragging more, suggesting either greater social incentives or better information by senior peers.

the extent of those working with a peer in a nurse-managed pod. Second, when a single physician is responsible for all patients entering the ED, I find no evidence of foot-dragging.

It is unlikely that these large effects can be explained by anything other than peer effects on foot-dragging. First, both regressions (6.1) and (6.2) include time dummies for hour of the day. More generally, the coefficients of interest in these regressions (and all other regressions identifying foot-dragging) correspond to responses of length of stay to expected future work, rather than levels of length of stay. Second, essentially all observations with only one physician in a pod occur during transition times of two to three hours during shifts in which the same physician works with a peer. The effect is thus identified by behavior of the same physician in the same shift, but only under different peer conditions. ED conditions are generally unchanged in this short window of time relative to nearby times.[38]

7 Patient Assignment and Equilibrium Building

I have shown that the self-managed system improves throughput productivity by reducing foot-dragging and that physician peers in the same pod observe more information about each other's true workloads. In this section, I test a prediction related to the *use* of this information by physicians in the self-managed system to assign patients. In the model in Section 2, following Milgrom and Roberts (1988), if physicians use more information about true relative workloads to choose patients in the self-managed system, foot-dragging should not only decrease, but the *ex post* assignment efficiency should also increase. That is, patients should be assigned according to censuses as signals of workload *more* often in the self-managed system than in the nurse-managed system.

In addition to patient assignment in steady state, a related but distinct issue is how physicians built the new equilibrium after Bravo pod switched from a nurse-managed to a self-managed system.[39] In Appendix Figure B-8.1, I show that foot-dragging does not immediately disappear after the switch in Bravo but rather takes a few months to disappear. Therefore, another question relates

[38] One obvious change when no peer is present is that workload per physician is greater. Results are qualitatively unchanged when normalizing measures of expected future work for the number of physicians in the ED.

[39] As discussed in Gibbons and Henderson (2012), equilibria may not be immediate or obvious, and they instead might need to be "built." They categorize reasons for this under needs to establish "credibility" and "clarity," and they review theoretical, experimental, and case-study literature on this (e.g., Greif, 1993; Fudenberg and Levine, 1993; Weber and Camerer, 2003). I discuss this issue – of why the equilibrium in Bravo after its regime change is not immediate – more specifically later in this section.

to how assignment patterns evolve and whether potentially foot-dragging physicians are more likely to get new patients during this transition period.[40]

7.1 Spatial Bed Patterns over Time

Before examining the relationship between censuses and assignment, I first examine spatial patterns in the use of beds to shed light on the transition between organizational systems in Bravo. A well-known experimental regularity is that players do not usually settle immediately on equilibria. Instead, people are influenced by the mere name of a game (Liberman et al., 2004) or irrelevant past experiences that could lead to "institutional afterglow" (Bohnet and Huck, 2004).

In other words, despite the fact that Bravo's organizational structure became isomorphic to Alpha's with the regime change, the new norm may not have been immediately "credible" or "clear" between peers (Gibbons and Henderson, 2012). Specifically, the regime change in Bravo was not announced as a move to replicate the environment in Alpha, but rather as a simple merger of bed ownership between peers with nothing else changed. Indeed, prescheduled Bravo shifts retained labels referring to obsolete divisions in Bravo (Bravo-1 and Bravo-2) until July 2010.

However, outside of experiments, little is known about the process of building equilibria. In this setting, the use of beds provides an insightful way of measuring behavior both in and transitioning to equilibria. In particular, I measure the extent to which the spatial patterns of beds used between physicians in Bravo continue to resemble obsolete nurse-managed zones of assignment even after Bravo's official switch to a self-managed system in March 2010. This exercise uses the fact that Bravo shifts retained obsolete zone labels until July 2010.

Figure 4 shows the distribution of patients according to bed location. Panel A shows the percentage of patients initially occupying a bed in Alpha for physicians working in Alpha. This percentage is close to 100% throughout the time series. Panel B shows the percentage of patients initially occupying a bed in Bravo-1. Percentages are shown for respective sets of physicians working in "Bravo-1" and "Bravo-2" labeled shifts, for as long as those labels exist until July 2010. Thereafter, percentages are shown for all Bravo physicians. Prior to the Bravo regime change, the median percentage of patients in Bravo-2 for physicians in Bravo-2 shifts is 0% throughout. The percentages for

[40] For more discussion of how patients are assigned in the self-managed system, see Section 2.3 and Appendix A-3. Regardless of whether physicians can commit to an assignment policy and of the amount of information they observe, patients should be assigned to physicians who foot-drag with greater probability if they have lower true workloads.

physicians in Bravo-1 shifts are substantially more varied, but the median percentage for physicians in these shifts hovers slightly above 90% until January 2010, when plans for the regime change is announced.[41] In March 2010, there is a sharp decrease in the Bravo-1 bed percentage for Bravo-1 physicians and a sharp increase in that percentage for Bravo-2 physicians. However, the move to the new self-managed equilibrium is not immediate. For the period during which Bravo shifts retain their obsolete "Bravo-1" and "Bravo-2" labels, until July 2010, physicians continue moving toward the new equilibrium.

7.2 Censuses and Assignment

Next, to examine the use of information between organizational systems, I study the relationship between censuses and patient assignment. Censuses are a public measure of workload, but they are signals that can be distorted by foot-dragging. A testable implication of the prediction of *ex post* assignment inefficiency is that physician censuses should be *less* related to new patient assignment in the nurse-managed system than in the self-managed system in equilibrium. Stated another way, this tests the theory that signals are used more efficiently in the self-managed system. In addition, assignment during the off-equilibrium transition of Bravo pod sheds light on how the equilibrium with no foot-dragging is eventually achieved in the self-managed system. Specifically, patient assignment to physicians who foot-drag should result in a positive correlation between assignment and censuses.

For my baseline specification, I estimate the linear probability model

$$Y_{ijt} = \alpha Census_{jt} + \beta \mathbf{ShiftTime}_{jt} + \gamma I_{jt \notin Self} ZoneLabel_{jt} + \eta_j + \nu_{it} + \varepsilon_{ijt}, \qquad (7.1)$$

where the outcome Y_{ijt} is an indicator variable that takes the value of 1 if patient i, arriving at the pod at time t, is assigned to physician j. $Census_{jt}$ denotes the number of patients under the care of physician j at time t and is the variable of interest. $\mathbf{ShiftTime}_{jt}$ is a vector of hourly time dummies of the duration between time t and the end of physician j's shift, since physicians

[41] Appendix B-7 provides more details about the locations and use of beds in the ED. Some of these details include the following: First, patients may occupy more than one bed and may occasionally be transferred between beds in different pods. Also, due to the physical layout of Bravo, Bravo-1 beds were less frequently used than those in Bravo-2. Finally, even in the nurse-managed system, physicians still must acknowledge patients, which provided some latitude for physicians to deviate from patient assignment strictly according to zones.

are less likely to be assigned new patients as they near the end of their shift, regardless of their censuses. For nurse-managed observations in Bravo (when $I_{jt \notin Self} = 1$), I also control for the zone (i.e., Bravo-1 or Bravo-2).[42] η_j controls for physician identities and allows for some physicians being more likely to take new patients regardless of census or observed behavior by their peers. ν_{it}, controlling for patient i at visit t, ensures that two physicians could have been assigned the patient. It implies that this linear probability model is equivalent to estimating a differenced model in which the variable of interest is $Census_{jt} - Census_{-jt}$, the difference in censuses between a physician and his peer, with a coefficient algebraically equal to α. The coefficient α represents the incremental likelihood, averaged over different shift times, with which a physician is to receive a new patient for each additional patient on his census relative to his peer's census.

Figure 5 shows a plot of coefficients α in (7.1) over time and in both pods.[43] I estimate α over each month by using triangular kernels with 45 days on each side of the first of the month; for months immediately before and after the regime change, I only use 45 days on the side away from the regime change. Prior to the Bravo regime change, Figure 5 shows relatively stable assignment in both pods. In both systems, physicians with lower censuses are more likely to be assigned patients, but this likelihood is consistently greater in the self-managed system (Alpha) than in the nurse-managed system (Bravo). Immediately after the regime change in March 2010, assignment in Bravo shows a jump in which physicians with *higher* censuses are actually more likely to receive new patients.[44] After three months, the spike reverses, and patients are again more likely to be assigned to physicians with lower censuses, even more so than prior to the regime change. Finally, Bravo's correlation between assignment and censuses after the spike settles to the same level as Alpha's.

These results show that, in equilibrium, the self-managed system improves the *ex post* assignment efficiency according to public signals of workload. By measuring assignment in both pods over time, I also show that assignment is not specific to pods, but rather to the organizational systems. This is consistent with the theory in Section 2: As the threat of foot-dragging is reduced in the self-managed system by the use of more information between peers, new patients can be more readily assigned to physicians with lower censuses. These results of the self-managed system are *not* consistent with

[42] Bravo-1 effectively has fewer beds than Bravo-2. In Appendix B-7, I discuss zone characteristics in greater detail.

[43] See Appendix Figures A-4.3 and A-4.4 for a set of plots representing the same estimates with confidence intervals.

[44] For robustness and given results in Section 7.1, I also modify Equation (7.1) to allow for residual zone norms after the Bravo regime change. Results are shown in Appendix Figure B-8.4. The positive spike is diminished but close to being significant at the 5% level.

an alternative mechanism, such as random patient assignment, that reduces foot-dragging but by ignoring all signals of workload.

Given that the self-managed equilibrium is not immediately established, and that residual foot-dragging remains after March 2010, a natural question is how full cooperation is eventually established. Again, some insight can be gained from the experimental literature. Enforcement in public goods games has been studied in seminal research by Ostrom et al. (1992) and Fehr and Gachter (2000). They have found that full cooperation is possible only when players are allowed, by the game's structure, to enforce it by punishment. Unlike laboratory experiments, this study cannot definitively show punishment. I do however see that, during the transitional period of residual foot-dragging in Bravo, physicians with higher censuses were *more* likely to be assigned new patients, reflected in the spike in Figure 5. This is consistent with foot-dragging physicians being made to take new patients in a newly self-managed system, possibly because their foot-dragging was observed by peers who thus decided not to take new patients. Such an assignment policy is consistent with eventually building a new equilibrium with no foot-dragging.

8 Conclusion

This paper studies an organizational change along a dimension akin to self-management, a key ingredient of teamwork, and finds a significant improvement in physician worker productivity. Simply by allowing physicians to choose patients, the self-managed system reduces patient lengths of stay by 11-15% relative to the nurse-managed system. This effect occurs primarily via the reduction of a moral hazard I call foot-dragging, in which physicians delay patient discharge to forestall new work. Foot-dragging is sensitive to peer effects, suggesting both social incentives and better information between peers in the same pod. The self-managed system additionally appears to use this information between peers further to improve patient assignment and reduce foot-dragging.

The size of the productivity increase is notable given the amount of information available through technology in this ED and the amount of private information in medical care that likely remains hidden even to peers in the same location. The intervention in this study is remarkably trivial in scope compared to other human resource management practices. In this sense, this study most likely provides an underestimate of the effect of teamwork as practiced in most settings, which may additionally involve other mechanisms. Nonetheless, this study isolates a key feature of teamwork – the joint management of work – that is likely to be important in a wide variety of settings. This feature is distinct from the classic concept of teams as joint production (e.g., Holmstrom, 1982). While the prediction of moral hazard is well known in joint production, teamwork reduces moral hazard by allowing agents, with better information than a more distant manager, to take part in the management of work.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller, "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program," *Journal of the American Statistical Association*, 2010, *105* (490), 493–505.

_ , _ , and _ , "Comparative politics and the synthetic control method," *MIT Political Science Department Research Paper*, 2011, *No. 2011-25.*

Acemoglu, Daron, Philippe Aghion, Claire Lelarge, John Van Reenen, and Fabrizio Zilibotti, "Technology, Information, and the Decentralization of the Firm," *The Quarterly Journal of Economics*, November 2007, *122* (4), 1759–1799.

Aghion, P. and J. Tirole, "Formal and real authority in organizations," *The Journal of Political Economy*, 1997, *105* (1), 1–29.

Alchian, Armen A. and Harold Demsetz, "Production, information costs, and economic organization," *The American Economic Review*, 1972, *62* (5), 777–795.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, "Social preferences and the response to incentives: Evidence from personnel data," *The Quarterly Journal of Economics*, 2005, *120* (3), 917–962.

_ , _ , and _ , "Social connections and incentives in the workplace: Evidence from personnel data," *Econometrica*, 2009, *77* (4), 1047–1094.

_ , _ , and _ , "Social incentives in the workplace," *The Review of Economic Studies*, 2010, *77* (2), 417–458.

Bertrand, Marianne and Antoinette Schoar, "Managing with Style: The Effect of Managers on Firm Policies," *The Quarterly Journal of Economics*, November 2003, *118* (4), 1169–1208.

_ , Esther Duflo, and Sendhil Mullainathan, "How Much Should We Trust Differences-In-Differences Estimates?," *The Quarterly Journal of Economics*, February 2004, *119* (1), 249–275.

Bliss, Christopher and Barry Nalebuff, "Dragon-Slaying and Ballroom Dancing: The Private Supply of a Public Good," *Journal of Public Economics*, November 1984, *25* (1), 1–12.

Bloom, Nicholas and John Van Reenen, "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics*, 2007, *122* (4), 1351–1408.

— , **Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts**, "Does Management Matter? Evidence from India," *The Quarterly Journal of Economics*, February 2013, *128* (1), 1–51.

Bohnet, Iris and Steffen Huck, "Repetition and Reputation: Implications for Trust and Trustworthiness When Institutions Change," *The American Economic Review*, 2004, *94* (2), 362–366.

Caroli, Eve and John Van Reenen, "Skill-Biased Organizational Change? Evidence from A Panel of British and French Establishments," *The Quarterly Journal of Economics*, November 2001, *116* (4), 1449–1492.

Crepon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora, "Do labor market policies have displacement effects? evidence from a clustered randomized experiment," *The Quarterly Journal of Economics*, May 2013, *128* (2), 531–580.

Delarue, Anne, Geert Van Hootegem, Stephen Procter, and Mark Burridge, "Teamworking and organizational performance: A review of survey-based research," *International Journal of Management Reviews*, 2008, *10* (2), 127–148.

Dessein, Wouter, "Authority and Communication in Organizations," *The Review of Economic Studies*, October 2002, *69* (4), 811–838.

Donnelly, Laura, "Minister: NHS will collapse if elderly bed blocking continues," *Telegraph.co.uk*, May 2013.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan, "Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India," *The Quarterly Journal of Economics*, November 2013, *128* (4), 1499–1545.

Elixhauser, Anne, Claudia Steiner, D. Robert Harris, and Rosanna M. Coffey, "Comorbidity Measures for Use with Administrative Data," *Medical Care*, January 1998, *36* (1), 8–27.

Fehr, Ernst and Simon Gachter, "Cooperation and punishment in public goods experiments," *The American Economic Review*, 2000, *90* (4), 980–994.

Forster, Alan J., Ian Stiell, George Wells, Alexander J. Lee, and Carl Van Walraven, "The Effect of Hospital Occupancy on Emergency Department Length of Stay and Patient Disposition," *Academic Emergency Medicine*, 2003, *10* (2), 127–133.

Fudenberg, Drew and David K. Levine, "Self-Confirming Equilibrium," *Econometrica*, May 1993, *61* (3), 523.

Gibbons, Robert and Rebecca Henderson, "What do managers do? Exploring persistent performance differences among seemingly similar enterprises," in Robert Gibbons and John Roberts, eds., *The Handbook of Organizational Economics*, Princeton, NJ: Princeton University Press, 2012, pp. 680–732.

Greif, Avner, "Contract enforceability and economic institutions in early trade: The Maghribi traders' coalition," *The American Economic Review*, 1993, *83* (3), 525–548.

Hamilton, Barton H., Jack A. Nickerson, and Hideo Owan, "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation," *Journal of Political Economy*, June 2003, *111* (3), 465–497.

Holmstrom, Bengt, "Moral Hazard in Teams," *The Bell Journal of Economics*, 1982, pp. 324–340.

Ichino, Andrea and Giovanni Maggi, "Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm," *The Quarterly Journal of Economics*, August 2000, *115* (3), 1057–1090.

Ichniowski, Casey and Kathryn Shaw, "Beyond incentive pay: Insiders' estimates of the value of complementary human resource management practices," *The Journal of Economic Perspectives*, January 2003, *17* (1), 155–180.

_ , _ , and Giovanna Prennushi, "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines," *American Economic Review*, 1997, *87* (3), 291–313.

Kandel, Eugene and Edward P. Lazear, "Peer Pressure and Partnerships," *The Journal of Political Economy*, 1992, pp. 801–817.

Knez, Marc and Duncan Simester, "Firm-wide Incentives and Mutual Monitoring at Continental Airlines," *Journal of Labor Economics*, October 2001, *19* (4), 743–772.

Lazear, Edward P., "Performance pay and productivity," *The American Economic Review*, December 2000, *90* (5), 1346–1361.

Lerman, Benjamin and Michael S. Kobernick, "Return Visits to the Emergency Department," *The Journal of Emergency Medicine*, September 1987, *5* (5), 359–362.

Liberman, Varda, Steven M. Samuels, and Lee Ross, "The Name of the Game: Predictive Power of Reputations versus Situational Labels in Determining Prisoner's Dilemma Game Moves," *Personality and Social Psychology Bulletin*, September 2004, *30* (9), 1175–1185. PMID: 15359020.

Manski, Charles F., "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 1993, *60* (3), 531–542.

Mas, Alexandre and Enrico Moretti, "Peers at Work," *The American Economic Review*, 2009, *99* (1), 112–145.

McHugh, Megan, Kevin Van Dyke, Mark McClelland, and Dina Moss, "Improving patient flow and reducing emergency department crowding: A guide for hospitals," Technical Report 11(12)-0094, Agency for Healthcare Research and Quality October 2011.

Meisel, Zachary and Jesse Pines, "Waiting Doom," *Slate*, June 2008.

Milgrom, Paul and John Roberts, "An Economic Approach to Influence Activities in Organizations," *The American Journal of Sociology*, 1988, *94* (Supplement), 154–179.

Nagin, Daniel S., James B. Rebitzer, Seth Sanders, and Lowell J. Taylor, "Monitoring, motivation, and management: the determinants of opportunistic behavior in a field experiment," *American Economic Review*, September 2002, *92* (4), 850–873.

Ostrom, Elinor, James Walker, and Roy Gardner, "Covenants With and Without a Sword: Self-Governance is Possible," *The American Political Science Review*, June 1992, *86* (2), 404.

Pallak, Michael S. and Robert O. Perloff, *Psychology and Work: Productivity, Change, and Employment.*, Washington DC: American Psychological Association, 1986.

Proudlove, N. C., K. Gordon, and R. Boaden, "Can good bed management solve the overcrowding in accident and emergency departments?," *Emergency Medicine Journal*, March 2003, *20* (2), 149–155. PMID: 12642528.

Rosenbaum, Paul, "Covariance adjustment in randomized experiments and observational studies," *Statistical Science*, August 2002, *17* (3), 286–304.

Roy, Donald, "Quota Restriction and Goldbricking in a Machine Shop," *American Journal of Sociology*, 1952, *57* (5), 427–442.

Rubin, Donald B., "Randomization analysis of experimental data: the fisher randomization test comment," *Journal of the American Statistical Association*, September 1980, *75* (371), 591–593.

Rubinstein, Ariel, "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 1982, *50* (1), 97–109.

Schuur, Jeremiah D. and Arjun K. Venkatesh, "The Growing Role of Emergency Departments in Hospital Admissions," *The New England Journal of Medicine*, 2012, *367*, 391–393.

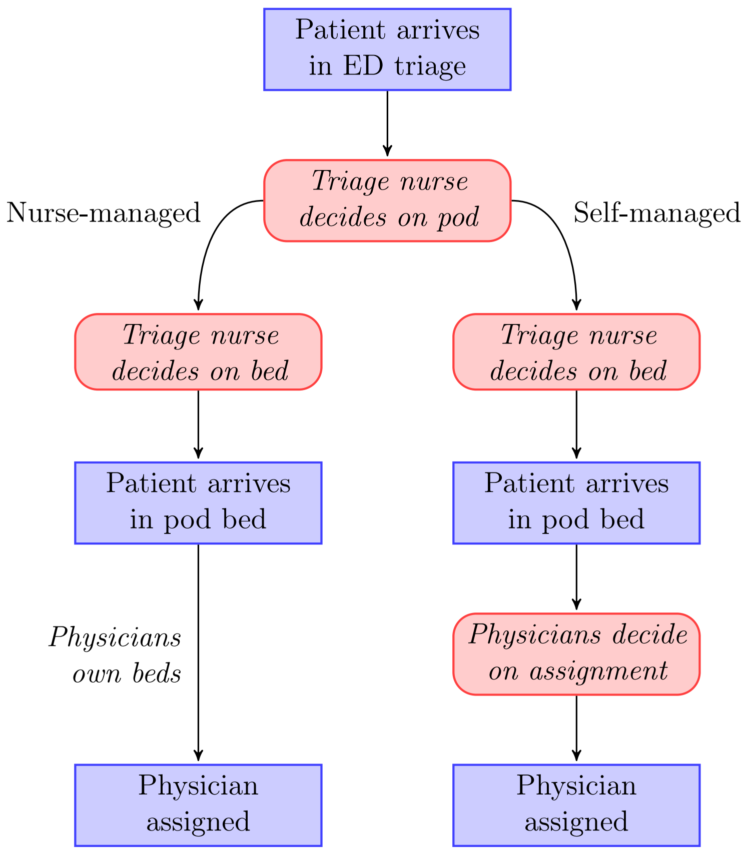
Syverson, Chad, "What Determines Productivity?," *Journal of Economic Literature*, June 2011, *49* (2), 326–365.

Tanabe, Paula, Rick Gimbel, Paul R. Yarnold, Demetrios N. Kyriacou, and James G. Adams, "Reliability and Validity of Scores on the Emergency Severity Index Version 3," *Academic Emergency Medicine*, 2004, *11* (1), 59–65.

Thompson, David A, Paul R Yarnold, Diana R Williams, and Stephen L Adams, "Effects of Actual Waiting Time, Perceived Waiting Time, Information Delivery, and Expressive Quality on Patient Satisfaction in the Emergency Department," *Annals of Emergency Medicine*, December 1996, *28* (6), 657–665.

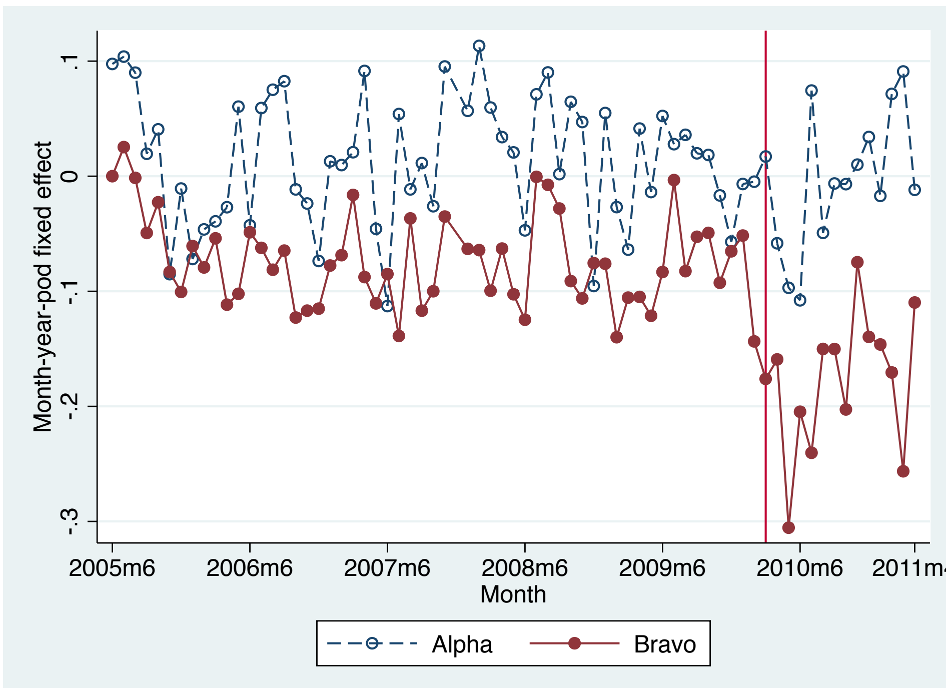
Weber, Roberto A. and Colin F. Camerer, "Cultural Conflict and Merger Failure: An Experimental Approach," *Management Science*, April 2003, *49* (4), 400–415.

Figure 1: Patient-to-physician Assignment Algorithm



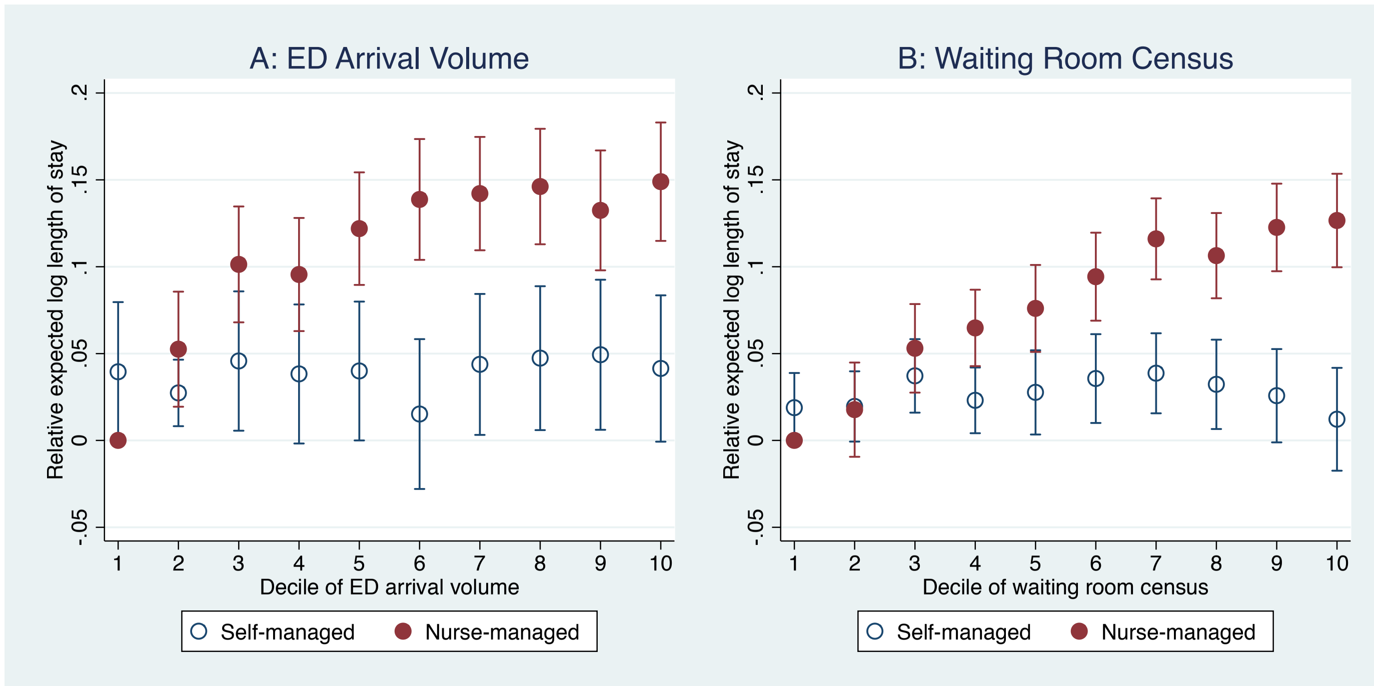
Note: This figure shows the patient assignment algorithm, starting with patient arrival at ED triage and ending with assignment to a physician. In ED triage, the triage nurse decides which pod and bed to send the patient. If the triage nurse decides to send the patient to a pod with a nurse-managed system (if one exists), then she also makes the decision on which physician will be assigned the patient because physicians own beds. If she decides to send the patient to a pod with a self-managed system, then she does not assign the physician. The physicians currently working in the self-managed pod will decide among themselves on that assignment. Although not shown in the figure, the triage nurse always assigns the bed and the nurse; she never assigns the resident, since residents in either pod choose their own patients or are told by physicians to see patients.

Figure 2: Overall Effect of Self-managed System on Log Length of Stay



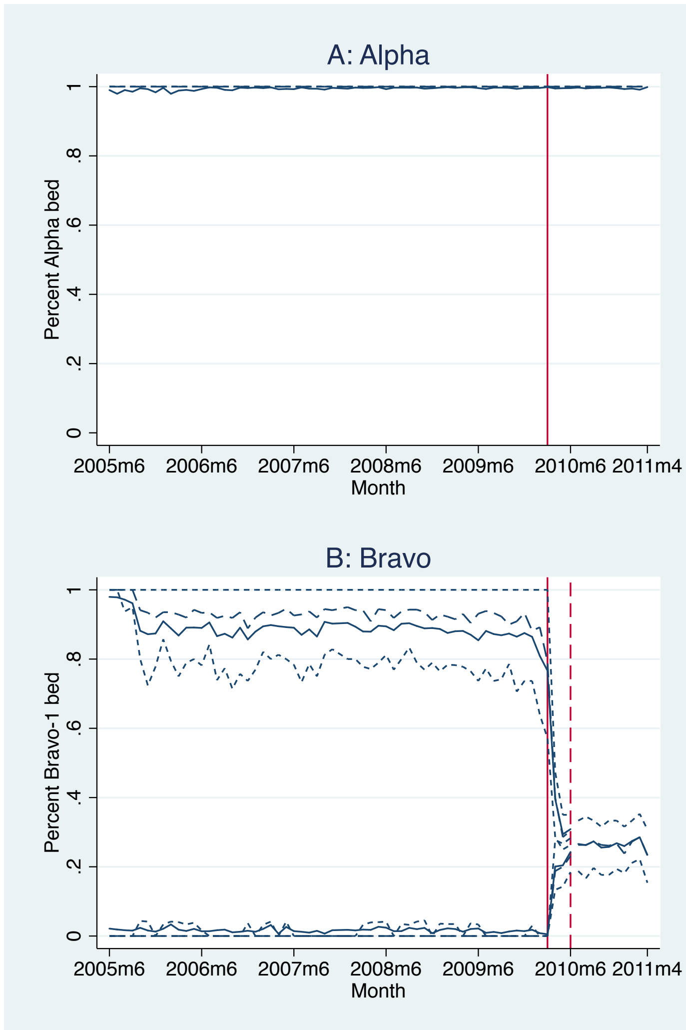
Note: This figure shows month-year-pod fixed effects estimated in a regression of log length of stay, as in Equation (4.2). Alpha pod fixed effects are plotted with hollow blue circles; Bravo pod fixed effects are plotted with solid red circles. The vertical red line indicates the month of the regime change of Bravo from a nurse-managed system to a self-managed system, in March 2010. Alpha was always self-managed. The fixed effect for Bravo in the first month is normalized to 0. The regression controls for uninteracted ED arrival volume, time categories, pod, patient demographics, patient clinical information, triage time, and physician-resident-nurse interactions.

Figure 3: Foot-dragging as Expected Future Work Increases by Deciles



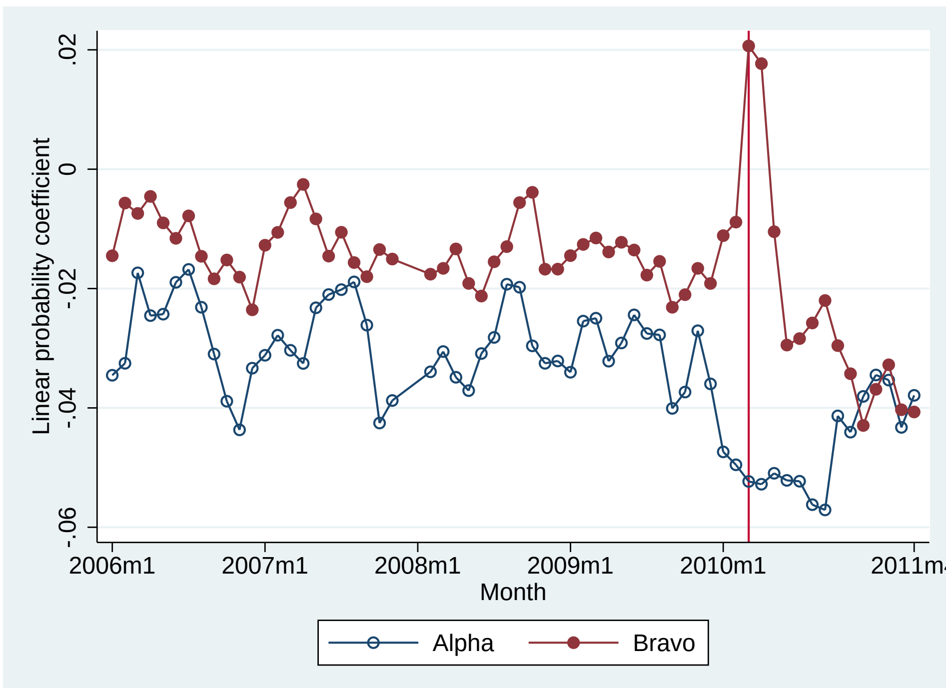
Note: This figure shows relative expected log length of stay as a function of expected future work. Panel A measures expected future work as ED arrival volume, or the number of patients arriving at triage in the hour prior to the patient's arrival at the pod. Panel B measures expected future work as waiting room census at the time of the patient's arrival at the pod. Expected log length of stay is normalized to 0 in the nurse-managed system and with the first decile of ED arrival volume. Coefficients for these decile-pod dummies are plotted from the regression of (5.2). Hollow blue circles indicate coefficients for the nurse-managed system; solid red circles indicate coefficients for the self-managed system. Vertical brackets show 95% confidence intervals.

Figure 4: Bed Location of Assigned Patients



Note: This figure shows bed locations of assigned patients for physicians working in Alpha and Bravo. Panel A shows the percentage of patients, of those under a physician with an Alpha shift, whose initial bed is recorded in Alpha pod. Panel B shows similar percentages over time for Bravo, but also considering two zones within Bravo: Bravo-1 and Bravo-2. Prior to the regime change in March 2010, physicians were assigned to work in either Bravo-1 or Bravo-2, and shifts were named accordingly. From March 2010, physicians in Bravo could see patients in either zone, but prescheduled shifts remained labeled with either "Bravo-1" or "Bravo-2" until the end of June 2010. Up to June 2010, two sets of lines are plotted in Panel B for the percentage of patients in initial Bravo-1 beds: the lower set for physicians with Bravo-1 shifts, the higher for those with Bravo-2 shifts. Weighted averages are plotted as solid lines, median percentages as long-dashed lines, and 25th and 75th percentiles of percentages as short-dashed lines. The March 2010 regime change is shown as a vertical solid red line; the ending of zone-specific Bravo shift labels in June 2010 is shown as a vertical dashed red line. Details of Bravo-1 and Bravo-2 bed locations are in Tables B-7.1 and B-7.3.

Figure 5: Effect of Additional Census Patient on New-patient Assignment Probability



Note: This figure shows the new-patient assignment probability, as a function of relative censuses for physicians within each pod. The plotted coefficient estimates from Equation 7.1 represent the average effect on assignment probability of each additional patient on a physician's census relative to his peer's census. Hollow blue circles show coefficient estimates for Alpha pod, which was always self-managed. Solid red dots show the coefficient estimates for Bravo pod, which switched to a self-managed system in March 2010, shown with a vertical red line. Coefficients are estimated in a kernel regression using a triangular kernel with 45 days on each side; estimates for February and March 2010 in Bravo pod are estimated by a kernel with 45 days only on the same side of the regime change. For simplicity, 95% confidence intervals are not plotted; see Appendix Figures A-4.3 and A-4.4 for plots with confidence intervals.

Table 1: Available Patients and ED Conditions for Physicians with Above- or Below-median Productivity

	Above-median productivity	Below-median productivity	t-statistic
Patient characteristic			
Age	48.7 (19.6)	48.7 (19.6)	-0.086
Emergency severity index	2.74 (0.78)	2.74 (0.78)	0.494
White	0.508 (0.500)	0.509 (0.500)	-0.038
Black or African-American	0.233 (0.423)	0.234 (0.423)	-0.069
Spanish speaking	0.098 (0.297)	0.097 (0.296)	0.203
Female and age $<$ 35 years	0.187 (0.390)	0.185 (0.388)	0.309
Prior ED patient volume			
Within last hour	6.06 (3.87)	5.97 (3.86)	0.917
Within last 6 hours	34.90 (19.11)	35.15 (18.95)	-0.418

Note: This table reports averages and standard deviations (in parentheses) for each characteristic of patients available to physicians with above-median and below-median productivity. "Available" means available to choose from in the self-managed system or assigned to in the nurse-managed system. Physician productivity is estimated by fixed effects in a regression of length of stay, controlling for team-member interactions, pod, patient characteristics, ED arrival volume, and time categories. The average difference in productivity between physicians of above- and below-median productivity is 0.28, meaning that physicians with above-average productivity have 28% shorter lengths of stay than those with below-average productivity. t-statistics for the difference in means are calculated assuming each shift is an independent observation and are all statistically insignificant at the 5% level.

Table 2: Overall Effect of Self-managed System on Log Length of Stay

	(1)	(2)	(3)	(4)	(5)
	Log length of stay				
Self-managed system	-0.108**	-0.113***	-0.120***	-0.133***	-0.148***
	(0.042)	(0.042)	(0.042)	(0.041)	(0.046)
ED arrival volume, time categories, pod dummies	Y	Y	Y	Y	Y
Patient demographics	N	Y	Y	Y	Y
Patient clinical information	N	N	Y	Y	Y
Patient triage time	N	N	N	Y	Y
Pod time trends	N	N	N	N	Y
Physician-resident-nurse dummies	Y	Y	Y	Y	Y
Number of observations	310,535	310,535	310,535	310,535	310,535
Adjusted R-squared	0.362	0.368	0.374	0.390	0.390
Sample mean log length of stay (log hours)	1.099	1.099	1.099	1.099	1.099

Note: This table reports the effect of the self-managed system on log length of stay, in Equation (4.1), while controlling for various observables. Column (5) is estimated with Equation (4.1), augmented with an additional term for pod time trends (i.e., $\gamma_p t$). All columns control for ED patient arrival volume, time categories (month-year, day of the week, and hour of the day dummies), pod, and team-member interactions. Various models may control for patient demographics; patient clinical information; and the time spent in triage, which reflects the triage nurse's subjective belief about patient severity. All models are also clustered by physician. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: Foot-dragging as Expected Future Work Increases

	(1)	(2)	(3)	(4)
	Log length of stay			
Measure of expected future work	ED arrival volume	ED arrival volume	Waiting census	Waiting census
Expected future work	0.006***	0.006***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Expected future work × self-managed system	-0.006***	-0.006***	-0.006***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Self-managed system	0.037	0.032	-0.004	-0.010
	(0.044)	(0.044)	(0.039)	(0.039)
Pod-specific volume control	N	Y	N	Y
Number of observations	282,105	282,105	282,105	282,105
Adjusted R-squared	0.439	0.440	0.439	0.440
Sample mean log length of stay (log hours)	1.179	1.179	1.181	1.181
Sample mean patient volume measure	15.58	15.58	8.78	8.78

Note: This table shows the effect of expected future work on log length of stay, estimated by Equation (5.1). Expected future work is measured either as the number of patients arriving at ED triage during the hour prior to the index patient's arrival at the pod ("ED arrival volume") or as the number of patients in the waiting room during that time ("waiting census"). Models (1) and (3) do not control for pod-level prior patient volume, defined as the number of patients arriving in the pod of the index patient one, three, and six hours prior to the index patient's arrival, while models (2) and (4) do. All models control for time categories (month-year, day of the week, and hour of the day dummies), pod, patient demographics (age, sex, race, and language), patient clinical information (Elixhauser comorbidity indices, emergency severity index), and physician-resident-nurse interactions. All models are clustered by physician. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: Effect of Peer Presence on Foot-dragging

	(1)	(2)	(3)
	Log length of stay		
Sample	Nurse-managed	Self-managed	Pooled
ED volume	0.006***	0.000	
	(0.001)	(0.001)	
ED volume × no peer present	0.024***	0.017***	0.018***
	(0.003)	(0.003)	(0.002)
ED volume × peer present, nurse-managed			0.009***
			(0.001)
ED volume × peer present, self-managed			-0.002**
			(0.001)
ED volume × only physician in ED			-0.002
			(0.002)
Number of observations	130,581	157,419	296,177
Adjusted R-squared	0.441	0.352	0.365
Sample mean log length of stay (log hours)	0.926	1.179	1.061
Sample mean ED volume	17.118	14.217	15.517

Note: This table reports the effect of expected future work, interacted with peer presence, on log lengths of stay. Expected future work is measured by ED arrival volume ("ED volume" for brevity), defined as the number of patients arriving at ED triage during the hour prior to the index patient's arrival at the pod. Equation (6.1) estimates models (1) and (2). Model (1) is estimated with observations of patients seen by nurse-managed teams; model (2) is estimated with self-managed teams. Both of these models use observations with at least one other physician in the ED, so that foot-dragging entails a negative externality against a current coworker, who may or may not be a peer. Model (3) is estimated by Equation (6.2) and includes the full sample. Main effects are included but omitted from the table for brevity. In all columns, the phrase "no peer present" means no other physician in same pod but another physician in ED, while "only physician in ED" means no other physician in entire ED. All models control for time categories, pod (when applicable), patient demographics, patient clinical information, triage time, and physician-resident-nurse interactions. All models are clustered by physician. * significant at 10%; ** significant at 5%; *** significant at 1%.