Teamwork and Moral Hazard: Evidence from the Emergency Department

David C. Chan

Stanford University

I investigate how teamwork may reduce moral hazard by joint monitoring and management. I study two organizational systems differing in the extent to which physicians may mutually manage work: Physicians are assigned patients in a "nurse-managed" system but divide patients between themselves in a "self-managed" system. The self-managed system increases throughput productivity by reducing a "foot-dragging" moral hazard, in which physicians prolong patient stays as expected future work increases. I find evidence that physicians in the same location have better information about each other and that, in the self-managed system, they use this information to assign patients.

I. 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, 127). Broad evidence suggests that teamwork and

I am 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, Stephen Ryan, Heidi Williams, a large number of seminar audiences, the editor (Jesse Shapiro), and anonymous referees. I gratefully acknowledge support from the National Bureau of Economic Research Health and Aging Fellowship, under the National Institute of Aging grant T32-AG000186; the Charles A. King Trust Postdoctoral Fellowship, the Medical Foundation; the Massachusetts Institute of Technology George and Obie Shultz Fund; and the Agency for Healthcare Research and Quality Ruth L.

Electronically published May 4, 2016
[*Journal of Political Economy*, 2016, vol. 124, no. 3]

other human resource management technologies are associated with higher productivity (Hamilton, Nickerson, and Owan 2003; Ichniowski and Shaw 2003; Bloom and Van Reenen 2007). 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 show theoretically that teamwork can reduce moral hazard by allowing workers to make use of better information about each other. I then study a natural experiment in which the same emergency department (ED) physicians work in two different organizational systems that differ only in the extent to which 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, "self-managed" system, the triage nurse assigns patients to a pod shared by two physicians, who then decide between themselves who will care for each arriving patient.

The natural experiment I study is well designed to permit me to estimate the effect of mutual management on productivity.[1] During the sample period, one pod operated under a self-managed system, while the other changed from a nurse-managed system to a self-managed system. Most providers work in both pods over time. Patient observations are frequent, with about one patient every 9 minutes, while the time period spanning the change is 6 years, allowing me to confirm the conditional parallel trends assumption and use systematic placebo tests for inference.

I find that physicians perform 11–15 percent 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 affects waiting times, a key determinant of patient satisfaction and health outcomes (e.g., Thompson et al. 1996), for patients to be seen. Although estimates for mortality and return visits are imprecise, I find no other difference in quality, financial, or process measures, suggesting that physicians simply delay discharging their patients and provide no more or less care for them.

While the effect of this treatment is intrinsically interesting, it is perhaps more important to understand the mechanisms behind such an effect. I develop a model that shows that, under asymmetric information between physician workers and the triage nurse manager in the nurse-managed system, physicians may want to avoid being assigned more work

Kirschstein Individual Postdoctoral Fellowship F32-HS021044. Data are provided as supplementary material online.

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

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 reducing foot-dragging relative to the nurse-managed system. However, other mechanisms could be at play. The self-managed system may improve outcomes through advantageous selection, as better patient-physician matches are formed according to either skill or availability. Alternatively, 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"). I therefore turn to the data for direct evidence on whether foot-dragging plays an important role in the effect of self-management.

I provide direct evidence for foot-dragging by showing that, controlling for the actual amount of current or future work, physicians are slower to handle current patients when the *expected* future workload is higher. This pattern holds only under nurse management and not under self-management, and the benefit of self-management disappears after accounting for this effect. I also find that the physical presence of a peer in the pod reduces foot-dragging in the nurse-managed system, consistent both with superior information (mutual monitoring) between peers and with social incentives against foot-dragging in the nurse-managed system.

Finally, I study patient assignment to test whether the self-managed system makes use of the better information between peers. I find that patient assignment is more negatively correlated with the number of patients currently in a physician's care in the self-managed system than in the nurse-managed system. This is consistent with the idea that under self-management the current workload becomes a less distorted picture of the physician's capacity for new patients, as in Milgrom and Roberts (1988). I also study patient assignment in the pod switching to a self-managed system and find evidence of enforcement against foot-dragging during an equilibrium-building transition period, in which physicians with *higher* censuses are more likely to be assigned new patients.

These findings are relevant to ED care, which costs $136 billion annually (Lee, Schuur, and Zink 2013), and to the growing health care sector of the US economy. More broadly, they suggest how firms may reduce moral hazard by allowing workers to manage each other. A large literature has shown wide variation in productivity across firms (Syverson 2011). High-level managers (Bertrand and Schoar 2003), management practices (Ich-

[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 (Proudlove, Gordon, and Boaden 2003; Meisel and Pines 2008), concerning enough for some observers to worry about the "collapse" of the UK National Health Service (Donnelly 2013).

niowski et al. 1997; Ichniowski and Shaw 2003; Bloom and Van Reenen 2007; Bloom et al. 2013), and group size (Gaynor, Rebitzer, and Taylor 2004) have been linked to productivity differences. However, key management features are difficult to describe and rarely isolated. This study adds to the literature by isolating an important feature of teamwork—mutual management—and illuminating a mechanism behind its effect.

Another strand of 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 that their peers can monitor them simply because they care about what their peers think (Kandel and Lazear 1992; Bandiera et al. 2005, 2009; Mas and Moretti 2009). 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 a distinct and important ingredient for workers to use shared information to improve productivity.[3]

The remainder of the paper proceeds as follows. Section II 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 III describes the ED institutional setting and data. Section IV reports the overall effect of the self-managed system. Sections V and VI discuss the main evidence for foot-dragging and its mitigation by organizational structure and the presence of peers. Section VII explores patient assignment in the two systems over time. Section VIII presents conclusions.

II. Conceptual 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 foot-dragging and improves assignment efficiency relative to the nurse-managed system, formalizing the concept that teamwork improves productivity by "monitoring and managing work processes" (Hackman 1986, 92).

It first is useful to clarify the focus of this model (and paper) in the context of related concepts. By joint production, I mean the process

[3] 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.

of more than one worker making an output (e.g., waiting times in this setting). I consider teamwork to be an organizational characteristic that allows workers involved in joint production to monitor and manage each other rather than to be managed by rules or an outside manager. Most workplaces allow some mutual monitoring, in which workers can observe each other better than an outside manager can (and possibly influence each other by social incentives; e.g., Mas and Moretti 2009). But teamwork additionally allows workers to manage each other's work. This study isolates mutual management as the key difference between the self-managed and nurse-managed systems and focuses on its role in reducing moral hazard.

The key feature of this model is that, by working together, workers may observe better information about each other's workload that they can then use to improve efficiency by teamwork. In the model, I assume that the triage nurse, as a distant manager, cannot observe true physician workloads, but that physician peers may sometimes observe true workloads. In the nurse-managed system, to forestall new work, physicians distort signals of true workload by prolonging patient lengths of stay (i.e., foot-dragging). Given this distortion, the triage nurse can be better off by committing to an ex post inefficient policy of ignoring workload signals, even though workload signals remain informative. In the self-managed system, however, physician peers may use information about each other's workload to assign new work. This reduces the threat of foot-dragging and improves ex post assignment efficiency.

A. *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 a low or a 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

$$u_j^P = -(t_j - \theta_j)^2 - K_P(\theta_j)\mathbb{I}\{\mathcal{J}(3) = j\}, \tag{1}$$

where t_j is the time that physician j keeps his initial patient, $\theta_j \in \{\underline{\theta}, \overline{\theta}\}$ is the workload entailed by his initial patient (where $\overline{\theta} > \underline{\theta} > 0$), $K_P(\theta_j) > 0$ is the cost of getting a potential third patient conditional on θ_j, and $\mathcal{J}(3)$ 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) is public information at any time. The action that each physician takes is t_j, which determines census $c_j \equiv \mathbb{I}\{t_a < t_j\}$, defined at the time of new patient arrival t_a. In the absence of 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).

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, given their preferences and financial incentives,[4] but that it is more costly for a physician with a high workload, for example, in terms of additional effort or worse outcomes for this new patient.

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.[5]
4. Exactly one patient will arrive with uniform probability distributed across the time interval $t_a \in [\underline{\theta}, \overline{\theta}]$. 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 (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.

B. Nurse-Managed System

In the nurse-managed system, the triage nurse assigns the new patient to a physician. I assume that physicians cannot report their types or any-

[4] 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 empirical predictions will be reversed; i.e., physicians having $\overline{\theta}$ will try to mimic having $\underline{\theta}$ by speeding up in order to get more work.

[5] I show in online app. A-2.1 that physicians have a dominant strategy in that it does not depend on the peer's type or strategy. Thus the order of stages 2 and 3 is unimportant.

thing else to the triage nurse but that the triage nurse can credibly commit to an assignment policy prior to physicians receiving their patients.[6]

The triage nurse's utility is

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

The variable 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 on. I specify $K_N(\underline{\theta}) = 0$ and $K_N(\overline{\theta}) = \overline{K}_N$, where $\overline{K}_N > 0$. This represents that it is managerially preferable to assign new work to a worker with a lower workload, for example, because that worker is able to handle the new work in a more timely or higher-quality manner. I do not restrict the value of $\overline{K}_N$ relative to $\overline{K}_P - \underline{K}_P$.[7]

The triage nurse commits to an assignment policy function $\pi_N(c_1, c_2)$, with censuses $c_j \in \{0, 1\}$, in which she assigns the new patient arriving at t to physician 1 with probability $\pi_N(c_1, c_2)$. To simplify the analysis, I impose a symmetric policy function with $\pi_N(0,0) = \pi_N(1,1) = \frac{1}{2}$ and $\pi_N \equiv \pi_N(0,1) = 1 - \pi_N(1,0)$, so that π_N can be understood as the probability that a physician with the lower census (if there is one) will be assigned the new patient. That is, aside from using information from physician censuses, the triage nurse's assignment policy treats the two physicians equally. Note that $\pi_N = 1$ represents what I mean by ex post efficiency, since if the triage nurse observes $c_j = 0$ and $c_{-j} = 1$, then she can infer that 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 $\pi_N^* \equiv \pi_N^*(0,1)$, given physician discharge strategies $\underline{t}^*$ and $\overline{t}^*$ for initial patients of type $\underline{\theta}$ and $\overline{\theta}$, respectively. Given π_N^*, physicians choose optimal discharge strategies $\underline{t}^*$ and $\overline{t}^*$.

Proposition 1. In the unique perfect Bayesian equilibrium for the nurse-managed system, physicians with $\underline{\theta}$ and $\overline{\theta}$ discharge their patients at $\underline{t}^* > \underline{\theta}$ (foot-dragging) 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 one-half and one (ex post inefficient assignment).

[6] In app. 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. Triage nurse commitment and physician reporting both increase efficiency.

[7] 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.

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 choose $\overline{t}^* = \overline{\theta}$. But physicians with $\underline{\theta}$ have some reason to mimic having $\overline{\theta}$ because the triage nurse would prefer to assign patients to a physician with a lower workload. For a given $\pi_N \equiv \pi_N(0,1)$ previously chosen by the triage nurse, the optimization problem of physicians with $\underline{\theta}$, $\max_{\underline{t}_j} \mathbb{E}[u_j^P(\underline{t}_j; \pi_N, \underline{\theta})]$, yields the following unique solution:[8]

$$\underline{t}^* = \underline{\theta} + \frac{\underline{K}_P}{2(\overline{\theta} - \theta)}\left(\pi_N - \frac{1}{2}\right). \tag{3}$$

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

The triage nurse commits to the assignment policy, π_N^*, that maximizes her expected utility given $\underline{t}^*$ and $\overline{t}^*$. Substituting (3) into her expected utility and solving the first-order condition yields π_N^*. For simplicity, I present π_N^* if $D = 0$:

$$\pi^*_{N|D=0} = \frac{1}{2} + \frac{(\overline{\theta} - \underline{\theta})^2}{\underline{K}_P}. \tag{4}$$

Equation (4) shows that the nurse's choice of π_N depends on the cost of getting the new patient for the low-workload physician, because of his temptation to foot-drag. Even if she wants to optimize only the assignment of the third patient (i.e., $D = 0$), committing to $\pi_N^* < 1$ (ex post inefficient assignment) may improve her expected utility, which is similar to Milgrom and Roberts's (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.[9]

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

[8] For convenience, I assume interior solutions throughout this analysis. Note that $\underline{t}^* \le \overline{\theta}$ and $\pi^* \le 1$. With respect to uniqueness, as discussed in more detail in app. A-2.1, 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 (i.e., physicians have a dominant strategy). This feature, combined with single solutions to both $\underline{t}^*$ and π_N^*, implies uniqueness of the perfect Bayesian equilibrium.

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

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 (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 $\Delta t \rightarrow 0$), physicians should expect to get a new patient the minute they discharge one. With no expected future patients (as $\Delta t \rightarrow \infty$), there is no incentive to foot-drag.

C. 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.

In this section, I focus on an assignment mechanism in which physicians commit to an assignment policy based on censuses, similar to the triage nurse's assignment policy, but with probability ψ may also use observations of true workload. An alternative mechanism is that physicians cannot commit to an assignment policy and can only choose a patient, playing a war of attrition (e.g., Bliss and Nalebuff 1984). Although both mechanisms likely contain some truth, I present the former mechanism in this section because playing a significant war of attrition is unlikely in the ED setting in which patients have urgent conditions. Note that an assignment policy also rules out the possibility of free riding, in which physicians delay accepting patients in the self-managed system in the hopes that their peer will accept the patient first.[10] I continue the baseline assumption that physicians cannot report their types. I present a brief discussion of results below; appendix A-3 contains details and derivation, as well as a more detailed treatment of the war-of-attrition mechanism and the possibility that physicians may report their types.

Proposition 2. Consider the unique perfect Bayesian equilibrium for the self-managed system. 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 ($\underline{t}^*$ will be lower and closer to $\underline{\theta}$) and more ex post efficient assignment ($\frac{1}{2} < \pi_N^* < \pi_S^* < 1$), relative to the nurse-managed system, and no free riding.

With probability ψ, physicians can use each other's observed types to assign the new patient, and only with probability $1 - \psi$, physicians must

[10] This is supported empirically in that physicians take only 20 minutes to write an order after a patient's bed arrival, which is an upper bound for the time that physicians take to sign up for patients. As discussed in Sec. IV and shown in table 3, I also find no evidence of an increase in time to write an order in the self-managed system.

assign the patient using a rule π_S based only on censuses. This lowers the attractiveness of foot-dragging, as the optimal strategy for a low-type physician is

$$\underline{t}^* = \underline{\theta} + (1-\psi)\frac{K_P}{2(\bar{\theta}-\underline{\theta})}\left(\pi_S - \frac{1}{2}\right).$$

Physicians can afford to commit to an assignment policy $\pi_S^* > \pi_N^*$, 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.[11]

The key point of this conceptual framework is that the threat of moral hazard is reduced in the self-managed system by mutual management that uses better information between peers. This framework deliberately abstracts from a few issues that would be useful to discuss here. First, this model abstracts from intrinsic worker heterogeneity: I rule out performance measurement (e.g., measuring average t_j) as a managerial tool to identify foot-dragging, because both workers are equally likely to foot-drag.

Second, it is natural to ask whether allowing physicians to report their workloads to the triage nurse can eliminate moral hazard. As shown in appendix A-2, physician reporting with discrete types eliminates foot-dragging, but foot-dragging is restored with continuous types. Consistent with the intuition in Myerson and Satterthwaite (1983), this reflects the fact that truth telling requires a restriction of types and messages that physicians can report. This restriction is especially unrealistic in the ED (and most medical contexts): Patient information is not only continuous but also multidimensional and complex, to the extent that communication problems exist even between doctors with no incentive misalignment (e.g., Apker, Mallak, and Gibson 2007).

Third, it is useful to contrast mutual management with social incentives, which also reduce moral hazard with mutual monitoring but through a social cost $S(\cdot)$ that peers incur when seen engaging in moral hazard (Kandel and Lazear 1992). As shown above, self-management can improve efficiency without social incentives, and social incentives likewise do not require mutual management (e.g., Mas and Moretti 2009). More formally, one could consider the interaction between organizational structure and social incentives by adding a term like $-\psi_G S_G(\cdot)$ into (1), in which components ψ_G and $S_G(\cdot)$ may depend on organizational structure

[11] In app. 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.

$G \in \{N, S\}$. While information between peers may endogenously increase with self-management (i.e., $\psi_S > \psi_N$), empirical work has also shown that social incentives may reverse signs depending on payoffs to peers. In particular, if the information asymmetry is primarily between peers and the triage nurse, peers in the self-managed system may jointly benefit from foot-dragging, encouraging the triage nurse to assign patients to another pod, while foot-dragging harms peers in the nurse-managed system. Social incentives could therefore encourage foot-dragging (i.e., "gold-bricking") in the self-managed system (Roy 1952; Bandiera et al. 2005).[12]

III. Institutional Setting and Data

I study a large, academic ED with a high frequency of patient visits with a total of 380,699 visits over 6 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, conditional on each assigned patient, to jointly produced waiting times. Patients are not shared and are rarely passed off between physicians. 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. 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., Lerman and Kobernick 1987; Forster et al. 2003; Schuur and Venkatesh 2012). Thirty-day mortality is perhaps the most unambiguous but occurs in only 2 percent of the sample. Hospital admission is a resource-intensive discharge option, which may substitute for appropriate care in the ED. Return visits from home within 14 days represent the complementary quality concern of premature home discharge.

[12] The possibility of collusion is not explicitly considered in this model since there are only two physicians, of whom one must be assigned the new patient; but this could be allowed in a model with a third physician in a different location who could be colluded against. Collusion of course reduces efficiency gains from self-management.

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.[13] 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 nursing, medication, laboratory, and radiology orders. I do not observe the time that a physician officially signs up for a patient, but I proxy for this using the time that the physician writes his first order.

A. *Organizational Systems*

Figure 1 illustrates the two organizational schemes that I study. 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.[14] 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 who assist physicians, that is, 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.[15] Financial incentives are also held constant: Physicians are paid a salary plus a 10 percent productivity bonus based on clinical productivity (measured by RVUs per hour).

Information about patients cared for by each physician is available to all physicians in the ED and the triage nurse from a computer interface (figs. B-8.1 and B-8.2). A salient summary statistic of workload is a physi-

[13] 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.

[14] The assignment of patients by nurses or nonmedical 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 or set financial incentives. In some settings, these nonmedical "managers" may have no discretion but merely follow rules.

[15] Unlike the intervention in Hamilton et al. (2003), the self-managed system involved no team incentives or expectations that physicians collaborate in patient care. More significant collaboration and other human resource management practices, including team incentives, could further improve productivity (e.g., Ichniowski et al. 1997; Hamilton et al. 2003), which I discuss in the conclusion.

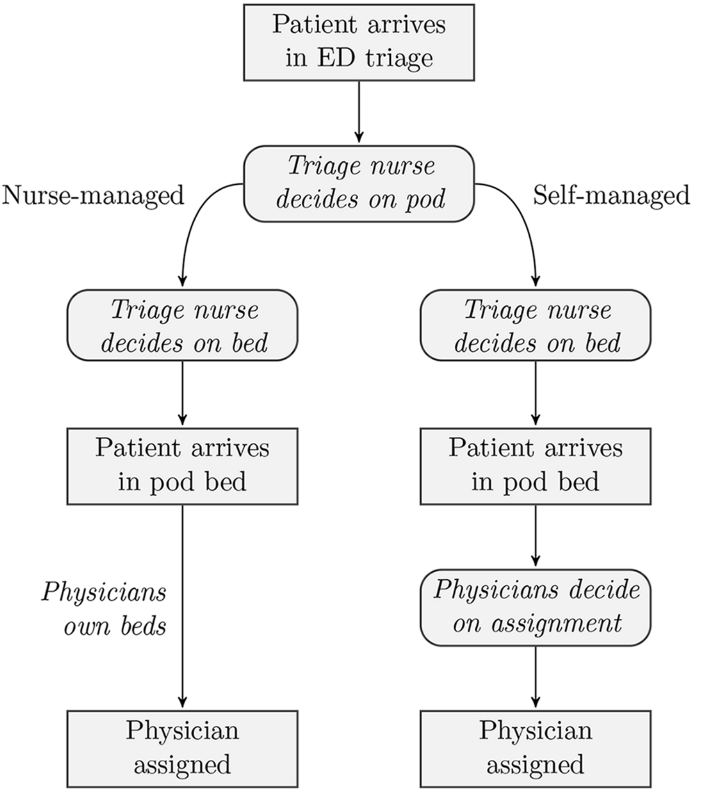


FIG. 1.—Patient-to-physician assignment algorithm. The figure shows the patient assignment algorithm, starting with patient arrival at ED triage and ending with assignment to a physician. In ED triage, the triage nurse decides which pod and bed to send the patient to. If the triage nurse decides to send 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.

cian's census, or the number of patients being cared for by the physician, but this of course can be distorted by foot-dragging. In both organizational systems, 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?).

B. Features for Identifying the Overall Effect

The ED has two pods: "Alpha" and "Bravo." Alpha pod has always had a self-managed system. Bravo pod used to be nurse managed but became self-managed in March 2010. 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 nonoverlapping 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 a 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.[16] 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. The switch involved no other administrative changes, including other interventions to promote teamwork.[17]

As an important time-invariant difference, Alpha pod has always been open 24 hours, while Bravo pod has always closed at night. As a result, patients who need to stay longer, because they are either sicker or have conditions that might make discharge difficult (e.g., psychiatric patients), 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.[18]

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, constituting 11,865 unique

[16] 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.

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

[18] 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 app. B-1.1. As I will discuss later, this works against finding improvements in productivity in Bravo.

physician-nurse-resident trios, are observed in both organizational systems.[19]

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, 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, 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).

C. *Features for Identifying Foot-Dragging*

Mutual management may affect productivity in multiple ways. In addition to reducing foot-dragging, workers may match to work better according to their skills and availability. If workers cannot commit to an assignment policy, they may also delay their acceptance of new work in the hopes that their peer will accept the work first ("free riding"). Finally, mutual management as a new task may incur an additional burden to workers and reduce their productivity.

My primary identification of foot-dragging relies on physician responses to shocks in expected future work, supported by two features of the setting. First, patient flow in the ED is highly unpredictable, even conditioning on rough time categories used for physician scheduling (app. B-5.1). Physician schedules are determined 1 year in advance, and physicians are able to request only rare specific shifts off, such as holidays or vacation days.[20] Once working on a shift, physicians cannot control the volume and types of patients arriving in the ED.[21] Conditional on the month-year, day of the week, and hour of the day, I find that physi-

[19] 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 the median resident is 1,525, while this number is 17 for residents who are observed to work in only one system.

[20] General preferences may be voluntarily stated in terms of rough time categories, but physicians are expected to be open for shifts at all times of the day and days of the week. Shift trades are also exceedingly rare, about less than one per month, or <1 percent of the number of shifts. Results are robust to dropping traded shifts. Per ED administration, shifts are not assigned with peers in mind.

[21] Physicians may rarely (<1–2 percent 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 affects only some ambulances (which as a whole constitute 15 percent of visits) carrying serious emergencies, as opposed to the majority of patients, some of whom walk in. ED flow is largely unaffected.

cians are exposed to similar patient types arriving at their pod and patient numbers arriving at the ED (table 1; app. B-6). This feature suggests that I can isolate shocks in expected future work that physicians cannot select or control.

Second, patients in the waiting room are yet to be assigned to pods. The distribution of patients between pods can vary widely. Depending on the time interval, the correlation between overall volume and pod-specific volume is at most .21, reflecting discretion by the triage nurse and stochastic discharge times creating differences in bed availability at the pod level (app. B-5.2). This variation allows me to separate expectations of future work (patients arriving at the ED but not yet assigned to a pod) from actual current or future work to a pod (patients who have been or will be assigned to the pod a physician is working in).

TABLE 1
Available Patients and ED Conditions for Physicians with Below- or Above-Median Productivity

	Below-Median Productivity	Above-Median Productivity	*t*-Statistic
Patient characteristic:			
Age	48.7	48.7	−.047
	(19.6)	(19.6)	
Emergency severity index	2.74	2.74	.451
	(.78)	(.78)	
White	.508	.509	−.050
	(.500)	(.500)	
Black or African American	.233	.234	−.068
	(.423)	(.423)	
Spanish-speaking	.098	.097	.212
	(.297)	(.296)	
Female and age <35 years	.187	.185	.332
	(.390)	(.388)	
Prior ED patient volume:			
Within last hour	6.06	5.97	1.076
	(3.87)	(3.86)	
Within last 6 hours	34.90	35.15	−.512
	(19.11)	(18.95)	

Note.—This table reports averages and standard deviations (in parentheses) for characteristics of available patients and ED patient volume that physicians with above-median and below-median productivity are exposed to. "Available patients" are patients assigned to the pod in the self-managed system and those assigned to the physician in the nurse-managed system. ED patient volume is represented by the number of patients arriving at the ED within the last hour or last 6 hours prior to an index patient's bed arrival. 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 percent shorter lengths of stay than those with below-average productivity. *t*-statistics for the difference in means are calculated assuming that each shift is an independent observation and are all statistically insignificant at the 5 percent level. More formal results on the quasi-random exposure of physicians to patient types and ED conditions are shown in online app. B-6.

IV. Overall Effect of the Self-Managed System

I estimate the overall effect of the self-managed system on a given team of providers for a given patient, asking the following: If the same patient and providers were in a different organizational system, what would their outcomes be? I 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 also control for providers because I observe the vast majority of providers—physicians, residents, and nurses—working in both pods over time.

I first show the effect of Bravo's regime change on length of stay graphically. Figure 2 shows month-year-pod fixed effects over time for the two pods estimated by this equation:

$$Y_{ijkpt} = \sum_{m=1}^{M}\sum_{y=1}^{Y} \alpha_{myp} I_{t\in m} I_{t\in y} + \beta \mathbf{X}_{it} + \tilde{\eta}\tilde{\mathbf{T}}_t + \nu_{jk} + \varepsilon_{ijkpt} \tag{5}$$

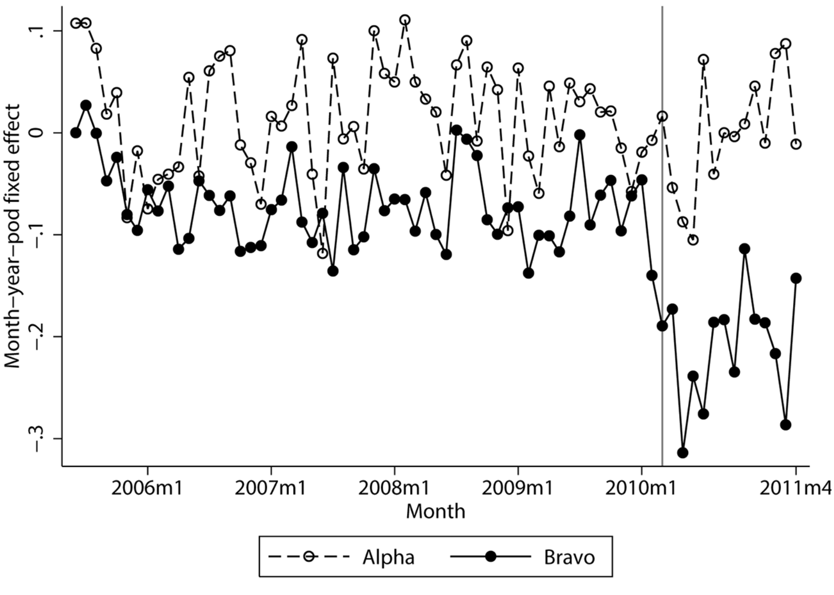


FIG. 2.—Overall effect of a self-managed system on log length of stay. This figure shows month-year-pod fixed effects estimated in a regression of log length of stay, as in equation (5). Alpha pod fixed effects are plotted with hollow circles; Bravo pod fixed effects are plotted with solid circles. The vertical gray 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 zero. The regression controls for noninteracted ED arrival volume, time categories, pod, patient demographics, patient clinical information, triage time, and physician-resident-nurse interactions.

for log length of stay Y_{ijkpt} indexed for patient i, physician j, resident-nurse k, pod p, and arrival time t. The parameters of interest are fixed effects for each month-year-pod, α_{myp}, where $I_{t\in m}$ and $I_{t\in y}$ are indicator functions for whether t belongs to month m and year y, respectively. I control for patient characteristics $\mathbf{X}_{it}$ (described in Sec. III.B), time categories $\tilde{\mathbf{T}}_t$ including the day of the week and hour of the day, pod identities ζ_p, and physician-resident-nurse trio identities ν_{jk}. The month-year-pod fixed effects capture average length of stay by location and time, after controlling for patient characteristics and worker identities. Two features of the fixed effects are apparent in figure 2. First, Bravo experienced a persistent decrease in length of stay after it changed to a self-managed system. Second, the two pods had roughly parallel trends in the long period prior to Bravo's regime change.

To quantify the self-managed effect on length of stay and other outcomes, I estimate the following equation:

$$Y_{ijkpt} = \alpha \mathrm{Self}_{pt} + \beta \mathbf{X}_{it} + \eta \mathbf{T}_t + \zeta_p + \nu_{jk} + \varepsilon_{ijkpt} \tag{6}$$

for outcome Y_{ijkpt}. The variable of interest in equation (6) is Self_{pt}, which indicates whether pod p had a self-managed system at time t. Time categories in the vector $\mathbf{T}_t$ include the month-year interaction, day of the week, and hour of the day.

Interpreting α in equation (6) as the causal effect of the self-managed system relies on the familiar assumption of conditionally parallel trends between Alpha and Bravo over time, which is supported by figure 2. In table 2, I estimate several versions of (6), 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 percent to -13 percent 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 (app. B-1). As there were no significant changes in peer characteristics between pods over time (app. B-1.2), adding controls for and interactions with peer characteristics does not change results (app. B-2). Adding pod-specific time trends (i.e., adding $\gamma_p t$) to equation (6) yields a slightly larger effect of -15 percent. Note finally that, since work can be sent to either pod, an improvement in throughput 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 also biases the estimated size of the causal effect downward (Rubin 1980).

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 equivalent to an in-

TABLE 2
Overall Effect of Self-Managed System on Log Length of Stay

	Log Length of Stay				
	(1)	(2)	(3)	(4)	(5)
Self-managed system	−.109**	−.114***	−.121***	−.133***	−.147***
	(.042)	(.042)	(.042)	(.041)	(.046)
ED arrival volume, time categories, pod dummies	Yes	Yes	Yes	Yes	Yes
Patient demographics	No	Yes	Yes	Yes	Yes
Patient clinical information	No	No	Yes	Yes	Yes
Patient triage time	No	No	No	Yes	Yes
Pod time trends	No	No	No	No	Yes
Physician-resident-nurse dummies	Yes	Yes	Yes	Yes	Yes
Observations	314,917	314,917	314,917	314,917	314,917
Adjusted R^2	.362	.368	.374	.390	.390
Sample mean log length of stay (log hours)	1.097	1.097	1.097	1.097	1.097

Note.—This table reports the effect of the self-managed system on log length of stay, in eq. (6), while controlling for various observables. Column 5 is estimated with eq. (6), 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. In app. B-4, I detail two alternative methods of inference, one assuming serially correlated pod-level error terms and another relying on systematic placebo tests. These yield p-values of .001 and .008, respectively.
* Significant at 10 percent.
** Significant at 5 percent.
*** Significant at 1 percent.

crease of 1.2–1.7 standard deviations in the productivity distribution across physicians.[22] 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.[23] 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 patient-level quality out-

[22] While the raw distribution of physician fixed effects has a standard deviation of 11 percent, distributions of physician effects adjusted by Bayesian shrinkage have standard deviations from 6.9 percent to 9.1 percent. Details are in app. B-3.

[23] 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.

comes (30-day mortality, hospital admissions, 14-day return visits) or financial/utilization outcomes (RVUs, total direct costs) at the 5 percent level, shown in table 3.[24] 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 significant effect on quality, revenue, and utilization between the two systems is more supportive of foot-dragging than other mechanisms such as better matching.

Any statement on quality, however, is limited by the lack of predictive power for mortality and return visits. While I can rule out a 0.7 percent increase in mortality, this is relatively large compared to the baseline 2.0 percent mortality. Nevertheless, estimates for admissions, RVUs, and costs are quite precisely estimated. For example, 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 −\$0.44 (95 percent confidence interval −\$2.32 to \$1.43). With no change in revenue or costs per patient, delaying the discharge of patients thus unambiguously decreases productivity from a financial perspective.

Table 3 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 zero and slightly negative.

An issue that arises in difference-in-differences estimation is the construction of appropriate standard errors for inference (Bertrand, Duflo, and Mullainathan 2004).[25] My baseline specification clusters standard errors by physician, suggesting an experiment sampling at the level of physicians, who are given shifts mapping to 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 of quasi-

[24] I focus on patient-level outcomes, rather than aggregate outcomes such as total RVUs per hour, for two reasons: (1) They readily allow adjustment for rich patient characteristics, which are important in this setting; and (2) they condition on patient arrival and assignment, which fluctuate nontrivially and especially as Bravo's equilibrium assignment of patients is still being established after its regime change, as shown in Sec. VII.

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

TABLE 3
Overall Effect of Self-Managed System on Other Outcomes

	30-Day Mortality (1)	Hospital Admissions (2)	14-Day Return Visits (3)	Relative Value Units (RVUs) (4)	Log Total Costs (5)	Log Time to First Order (6)
Self-managed system	.0002	.0004	−.0127*	−.013	−.016	−.025
	(.0031)	(.0086)	(.0067)	(.028)	(.030)	(.024)
Observations	317,199	317,199	317,199	255,516	284,965	307,885
Adjusted R^2	.303	.463	−.025	.405	.529	.126
Sample mean outcome	.020	.272	.067	2.701	6.724	−.623

Note.—This table reports the effect of the self-managed system on outcomes other than length of stay, estimated by eq. (6). RVUs represent intensity of care and are directly scalable to dollar amounts of clinical revenue. Total costs include direct costs for the entire visit, which may include hospital admission. Time to first order is the time between patient arrival at the pod and the first physician order, measured in log hours. All models control for ED patient arrival volume, time categories (hour of the day, day of the week, and month-year dummies), pod, patient demographics (age, sex, race, and language), patient clinical information (Elixhauser comorbidity indices, emergency severity index), triage time, and physician-resident-nurse interactions. These controls correspond to col. 4 of table 2, although results are robust to other controls (shown in app. table B-9.1). All models are clustered by physician.
* Significant at 10 percent.
** Significant at 5 percent.
*** Significant at 1 percent.

random exposure to patients and peers, conditional on rough time categories (app. B-6).

Still, to address the additional statistical issue of unobserved and potentially correlated pod-level shocks over time, I consider two alternative approaches that exploit the long time-series dimension shown in figure 2. First, I parametrically model sampling variation by assuming a month-year-pod shock correlated by a first-order autoregressive process across months within a pod. Second, I compare my observed effect (in Bravo after March 2010) with placebo regime changes in both pods over many months. In the spirit of systematic placebo tests (Abadie, Diamond, and Hainmueller 2010) and randomization inference (Rosenbaum 2002), this latter approach makes no assumption on sampling variation but simply asks how unusual the effect is relative to those obtained for the placebo regime changes. Detailed in appendix B-4, both approaches yield a high degree of statistical significance, with p-values of .001 and .008, respectively.

V. 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.

I first show graphical evidence of increasing length of stay with expected future work. Figure 3 plots coefficients $\{\alpha_0^2, \ldots, \alpha_0^{10}; \alpha_1^1, \ldots, \alpha_1^{10}\}$ for each decile d of expected future work, EDWork_t, interacted with organizational system in the equation

$$\begin{aligned} Y_{ijkpt} = & \sum_{d=2}^{10} \alpha_0^d (1 - \text{Self}_{pt}) D_d(\text{EDWork}_t) \\ & + \sum_{d=1}^{10} \alpha_1^d \text{Self}_{pt} D_d(\text{EDWork}_t) + \beta \mathbf{X}_{it} + \eta \mathbf{T}_t \\ & + \zeta_p + \nu_{jk} + \varepsilon_{ijkpt} \end{aligned} \tag{7}$$

for log length of stay Y_{ijkpt} for patient i, physician j, resident-nurse k, pod p, and time t. The variable $D_d(\text{EDWork}_t)$ equals one if expected future work in the ED at time t, EDWork_t, is in the dth decile. As before, Self_{pt} indicates whether pod p at time t was self-managed, and I control for patient characteristics $\mathbf{X}_{it}$, time categories $\mathbf{T}_t$ (month-year interaction, day of the week, hour of the day), pod fixed effects ζ_p, and provider-trio fixed effects ν_{jk}. The coefficients $\{\alpha_0^2, \ldots, \alpha_0^{10}; \alpha_1^1, \ldots, \alpha_1^{10}\}$ can be interpreted as the relative expected length of stay for patients in different organizational systems and under different states of 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 zero.

I measure expected future work, EDWork_t, in two ways. First, I consider ED arrival volume, or 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 not yet known. Second, I consider 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 waiting room census, this is a more salient measure of expected future work since physicians can readily click on the computer interface to see this census. By $\mathbf{T}_t$, I control for unobserved patient and ED characteristics correlated with both time categories and expected future work. As shown in figure 3, lengths of stay progressively increase in the nurse-managed system as expected future work increases, using both measures, consistent with the intuition that the incentive to foot-drag continues to

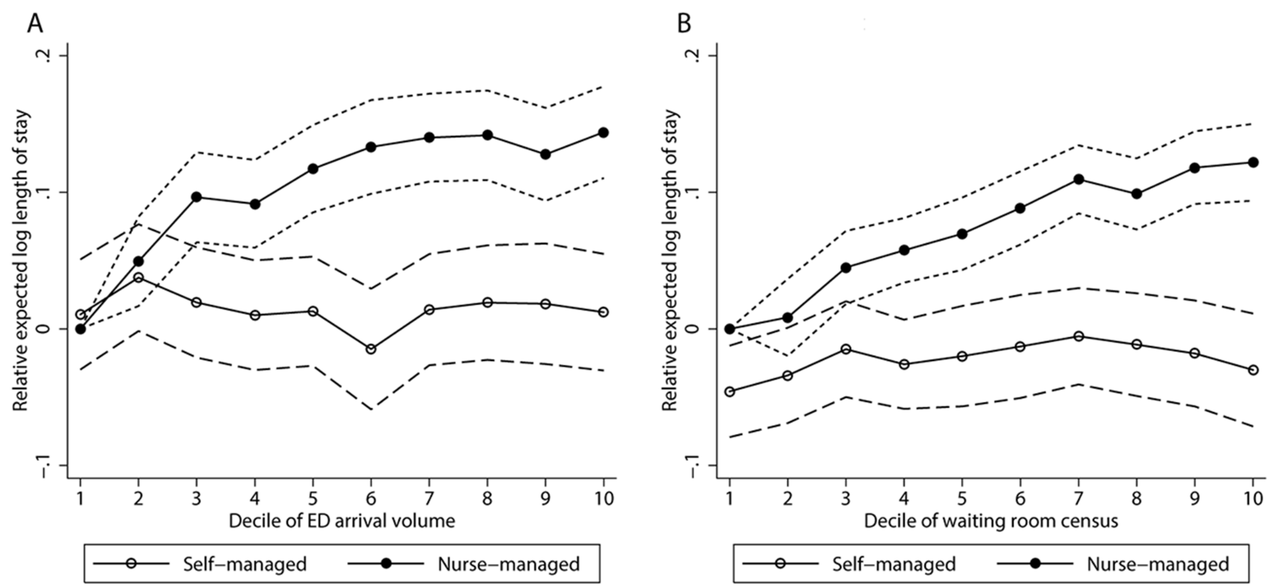


FIG. 3.—Foot-dragging as expected future work increases. The figure shows relative expected log length of stay as a function of expected future work. Panel *A* measures expected future work as ED arrival volume, or the number of patients arriving at triage in the hour prior to the patient's arrival at the pod. Panel *B* measures expected future work as the waiting room census at the time of the patient's arrival at the pod. Expected log length of stay is normalized to zero in the nurse-managed system and with the first decile of ED arrival volume. Coefficients for these decile-pod dummies are plotted from estimates of equation (7). Hollow circles indicate coefficients for the nurse-managed system; solid circles indicate coefficients for the self-managed system. Long- and short-dashed lines show 95 percent confidence intervals for the nurse-managed and self-managed coefficients, respectively.

grow as expected future work increases.[26] The self-managed system does not show increases in length of stay with patient volume and has roughly the same length of stay as the nurse-managed team at low patient volumes.

For the magnitude of foot-dragging, I estimate equations of the following form:

$$\begin{aligned} Y_{ijkpt} = {} & \alpha_1 \mathrm{EDWork}_t + \alpha_2 \mathrm{Self}_{pt} \cdot \mathrm{EDWork}_t + \alpha_3 \mathrm{Self}_{pt} \\ & + \beta \mathbf{X}_{it} + \eta \mathbf{T}_t + \zeta_p + \nu_{jk} + \varepsilon_{ijkpt}. \end{aligned} \tag{8}$$

The coefficients of interest in (8) 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 4 reports estimates for (8) 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 percent 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.[27] The coefficient for the interaction between expected future work and 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 −1.3 percent to 3.6 percent. This suggests that, in the limit with no expected future work, the two systems produce roughly similar lengths of stay. Results are robust to including controls for future or current pod-level work in (8). This suggests that physicians are responding differently because of expectations of future work rather than handling work assigned to their pod differently (e.g., free riding or advantageous selection) between the two organizational systems.

My measures of expected future work are likely to be noisy representations of physicians' true expectations of future work. Therefore, this

[26] Section II predicts strictly increasing foot-dragging with expected future work because physicians always have availability to receive a new patient. This is consistent with practice, given the large number of locations in which to place patients, including hallway spaces, and the need to keep beds open for especially emergent visits. Bed locations are detailed in app. B-8 and tables B-8.1–B-8.4.

[27] I estimate this by using log measures of expected future work as EDWork_t. My preferred specification, shown in table 4, does not take logs of expected future work because it is roughly normally distributed. However, results are qualitatively the same in this specification.

TABLE 4
Foot-Dragging as Expected Future Work Increases

	Log Length of Stay			
	Measure of Expected Future Work Is ED Arrival Volume		Measure of Expected Future Work Is Waiting Census	
	(1)	(2)	(3)	(4)
Expected future work	.0064***	.0055***	.0060***	.0063***
	(.001)	(.0009)	(.0014)	(.0013)
Expected future work × self-managed system	−.0059***	−.0059***	−.0060***	−.0060***
	(.0015)	(.0014)	(.0015)	(.0015)
Self-managed system	.0356	.0290	−.0051	−.0131
	(.0436)	(.0444)	(.0387)	(.0396)
Pod-specific volume control	No	Yes	No	Yes
Observations	286,275	286,275	286,275	286,275
Adjusted R^2	.439	.440	.439	.440
Sample mean log length of stay (log hours)	1.182	1.182	1.182	.182
Sample mean patient volume measure	15.57	15.57	8.78	8.78

Note.—This table shows the effect of expected future work on log length of stay, estimated by eq. (8). 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 1, 3, and 6 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), patient triage time, and physician-resident-nurse interactions. Controls are equivalent to controls in col. 3 of table 2, although results are robust to including other sets of controls in table 2. All models are clustered by physician.
* Significant at 10 percent.
** Significant at 5 percent.
*** Significant at 1 percent.

estimate of foot-dragging 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 be 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. Since length of stay does not increase with expected future work in the self-managed system, foot-dragging relative to zero and foot-dragging relative to self-managed teams are roughly the same in magnitude.

I also estimate the baseline equation for foot-dragging, equation (8), for other outcomes and process measures: 30-day mortality, admissions, 14-day return visits, RVUs, total direct costs, and a host of detailed process measures including laboratory, medication, and radiology orders. As shown in appendix table B-9.2, 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 decreases in hospital admissions and total costs, including 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.

VI. Peer Effects on Foot-Dragging

A key requirement for teamwork to reduce moral hazard is that workers may engage in mutual monitoring and therefore have better information about each other's workload than a manager might have. In this section, I show evidence that physicians working together observe better information about each other's workloads. While physicians usually work in pods with a peer, during certain times on a 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.[28]

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.[29] 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).

[28] 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 app. B-7.1, where I show results similar to those of Mas and Moretti (2009) (i.e., productive peers increase the productivity of physicians).

[29] As described in Sec. VII.B and app. B-8, physicians in the nurse-managed system are responsible for a fixed number of beds, regardless of peer presence. Even in the self-managed system, where beds are not owned, average censuses for physicians are similar during times when physicians are alone in a pod vs. with a peer.

Second, when there is only one physician in the entire ED, 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.[30]

In table 5, I present results for regressions of the form

$$\begin{aligned} Y_{ijkpt} = {} & \alpha_1 \mathrm{EDWork}_t + \alpha_2 \mathrm{NoPeer}_{jt} \cdot \mathrm{EDWork}_t + \alpha_3 \mathrm{NoPeer}_{jt} \\ & + \beta \mathbf{X}_{it} + \eta \mathbf{T}_t + \zeta_p + \nu_{jk} + \varepsilon_{ijkpt} \end{aligned} \tag{9}$$

for nurse-managed-team and self-managed-team samples separately. The variable 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 equation (9) only when there are at least two physicians in the ED, so that foot-dragging always entails a negative externality against a current coworker. I control for time categories, $\mathbf{T}_t$, and thus unobserved patient types and ED conditions during times of the day when physicians are likely to be alone.

I also perform the pooled regression

$$\begin{aligned} Y_{ijkpt} = {} & \sum_{s=1}^{4} 1(\mathrm{PeerState}_{jt} = s)(\alpha_s \mathrm{EDWork}_t + \delta_s) \\ & + \beta \mathbf{X}_{it} + \eta \mathbf{T}_t + \zeta_p + \nu_{jk} + \varepsilon_{ijkpt}, \end{aligned} \tag{10}$$

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 5 are consistent with previously estimated coefficients for the increase in length of stay with expected future work, shown in table 4, 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 columns 1 and 2 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 (10), shown in column 3 of table 5, confirm this and show that physicians do

[30] 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 bias measured foot-dragging upward only if the omitted volume of actual work is positively correlated with expected future work and increases length of stay.

TABLE 5
Effect of Peer Presence on Foot-dragging

	Log Length of Stay		
	Nurse-Managed Sample (1)	Self-Managed Sample (2)	Pooled Sample (3)
ED volume	.0059***	.0003	
	(.0012)	(.0019)	
× no peer present	.0253***	.0159**	.0178***
	(.0043)	(.0063)	(.0031)
× peer present, nurse-managed			.0089***
			(.0012)
× peer present, self-managed			−.0022
			(.0020)
× only physician in ED			−.0018
			(.0027)
Observations	130,029	165,109	295,060
Adjusted R^2	.442	.353	.365
Sample mean log length of stay (log hours)	.926	1.168	1.061
Sample mean ED volume	17.118	14.256	15.520

Note.—This table reports the effect of expected future work, interacted with peer presence, on log length of stay. Expected future work is measured by ED arrival volume (ED volume for short), defined as the number of patients arriving at ED triage during the hour prior to the index patient's arrival at the pod. Equation (9) 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 eq. (10) 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 the same pod but another physician in ED, while "only physician in ED" means no other physician in the entire ED. As in table 4, 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 percent.

** Significant at 5 percent.

*** Significant at 1 percent.

not increase lengths of stay with expected future work when they are alone in the ED.[31]

These results suggest that physicians reduce foot-dragging when a peer is present, consistent with peers observing true workload better than physicians in a different location and with social incentives.[32] Addi-

[31] These results are robust to considering expected future work as ED arrival volume per physician present, shown in table B-9.3, although the increase in foot-dragging with no peer present is smaller.

[32] In app. B-7.2, I show that peer effects on foot-dragging can also depend on the relationship with the peer present. In particular, working with senior peers (those having started employment at least 2 years earlier) diminishes foot-dragging more, suggesting either greater social incentives or better information by senior peers.

tionally, the two falsification tests, in columns 2 and 3 of table 5, 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 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.

VII. 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 examine a prediction related to the use of this information by physicians in the self-managed system to assign patients. In the model in Section II, following Milgrom and Roberts (1988), if physicians use more information about true relative workloads to choose patients in the self-managed system, not only should foot-dragging decrease, but the ex post assignment efficiency should also increase ($\pi_S^* > \pi_N^*$). That is, the use of better information between workers allows signals to be less subject to distortion and ultimately more valuable for making decisions. I test this prediction by measuring whether new patient assignment incorporates information in existing censuses, as signals of workload, more often in the self-managed system than in the nurse-managed system.

In addition to patient assignment in a 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. As discussed in Gibbons and Henderson (2012), equilibria may not be immediate or obvious, and they instead might need to be "built."[33] In appendix figure B-9.1, I show that foot-dragging does not immediately disappear after the switch in Bravo but rather takes a few months to disappear. Examining behavior off equilibrium provides a unique opportunity to study the mechanisms that enforce equilibria, distinguishing self-management (via patient assignment) from social incentives, for example.[34] Therefore,

[33] Gibbons and Henderson (2012) categorize reasons for this under needs to establish "credibility" and "clarity," and they review theoretical, experimental, and case study literature on this (e.g., Fudenberg and Levine 1993; Greif 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.

[34] Off equilibrium, under self-management, patients should be assigned with greater probability to physicians who foot-drag, regardless of whether physicians can commit to an assignment policy and of the amount of information they observe. Under social incentives, foot-dragging physicians would incur a social cost but would otherwise not be assigned more patients. For more discussion of how patients are assigned in the self-managed system, see Sec. II.C and app. A-3.

another question relates to how assignment patterns evolve and whether potentially foot-dragging physicians are more likely to get new patients during this transition period.

A. *Censuses and Assignment*

Next, to examine the use of information between organizational systems, I study the relationship between censuses and patient assignment. For my baseline specification, I estimate the linear probability model

$$\begin{aligned} Y_{ijt} = {} & \alpha \text{Census}_{jt} + \beta \mathbf{ShiftTime}_{jt} + \gamma I_{jt \notin \text{Self}} \text{ZoneLabel}_{jt} \\ & + \eta_j + \nu_{it} + \varepsilon_{ijt}, \end{aligned} \tag{11}$$

where the outcome Y_{ijt} is an indicator variable that takes the value of one 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 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 $I_{jt \notin \text{Self}} = 1$), I also control for the zone (i.e., Bravo 1 or Bravo 2).[35] The variable η_j controls for physician identities and allows for some physicians to be more likely to take new patients regardless of census or observed behavior by their peers. Controlling for patient i at visit t, ν_{it} 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 regressor of interest is $\text{Census}_{jt} - \text{Census}_{-jt}$, 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 4 shows a plot of coefficients α in (11) over time and in both pods. I estimate α over each month with a local linear regression, using triangular kernels with 90 days on each side of the first of the month, truncated if sufficiently close to March 1, 2010. Prior to the Bravo regime change, figure 4 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 af-

[35] Bravo 1 effectively has fewer beds than Bravo 2. In app. B-8, I discuss zone characteristics in greater detail.

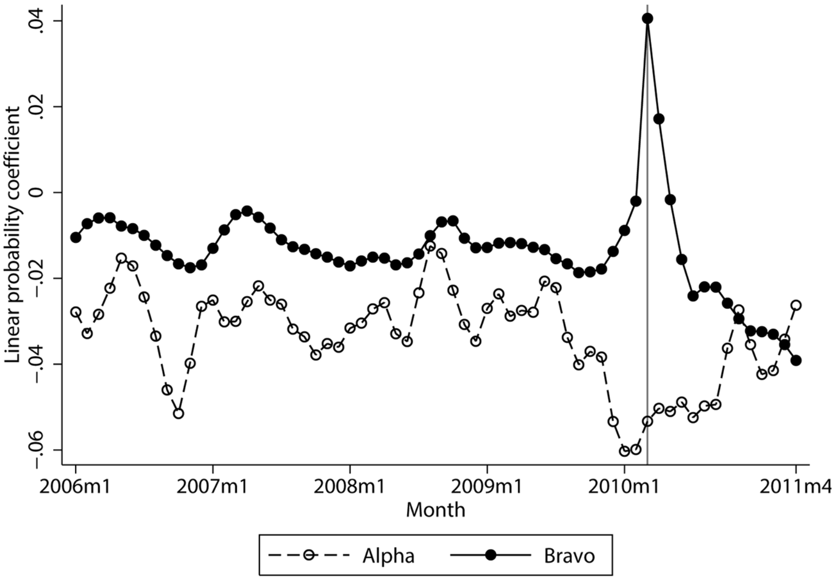


FIG. 4.—Effect of an additional census patient on the new patient assignment probability. The figure shows the new patient assignment probability, as a function of relative censuses for physicians within each pod. The plotted coefficient estimates from equation (11) represent the average effect on assignment probability of each additional patient on a physician's census relative to his peer's census. Hollow circles show coefficient estimates for Alpha pod, which was always self-managed. Solid circles show the coefficient estimates for Bravo pod, which switched to a self-managed system in March 2010, shown with a vertical gray line. Coefficients are estimated in a local linear regression using a triangular kernel with 90 days on each side; estimates for February and March 2010 in Bravo pod are estimated by a kernel with 45 days, which may be truncated if sufficiently close to March 1, 2010, on either side. For simplicity, 95 percent confidence intervals are not shown; see appendix figure B-9.2 for plots with confidence intervals. Appendix figures B-9.3 and B-9.4 show alternative specifications.

ter 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.[36] After 3 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

[36] For robustness and given results in Sec. VII.B, I also modify eq. (11) to allow for residual zone norms after the Bravo regime change. Results are shown in fig. B-9.3. The positive spike is diminished but is still significant at the 5 percent level. Results are also largely unchanged using a kernel regression, as shown in fig. B-9.4.

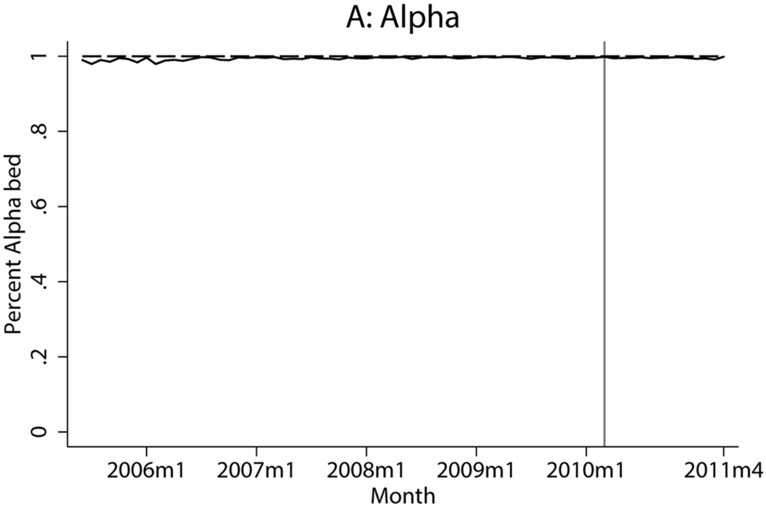
workload. By measuring assignment in both pods over time, I also show that assignment is not specific to pods, but rather to the organizational systems. This is consistent with the theory in Section II: As the threat of foot-dragging is reduced in the self-managed system by the use of more information 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 an alternative mechanism that reduces foot-dragging but by ignoring all signals of workload (e.g., random patient assignment).

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, Walker, and Gardner (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 4. 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.

B. *Spatial Bed Patterns over Time*

I provide more direct evidence of equilibrium building by examining how spatial patterns of bed assignment evolve over time. 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. Using this fact, 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.

Figure 5 shows the distribution of patients according to bed location. Panel *A* shows the percentage of patients initially occupying a bed in



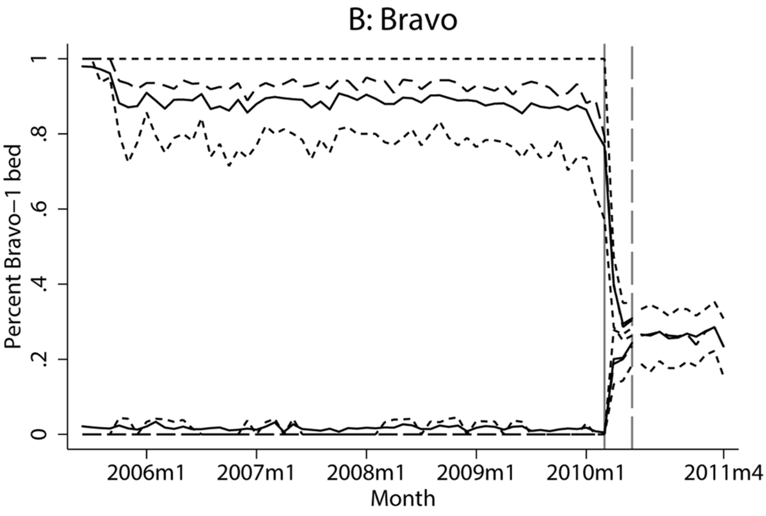


Fig. 5.—Bed location of assigned patients. The figure shows bed locations of assigned patients for physicians working in Alpha and Bravo. Panel *A* shows the percentage of patients, of those under a physician with an Alpha shift, whose initial bed is recorded in Alpha pod. Panel *B* shows similar percentages 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 upper set for physicians with Bravo 1 shifts and the lower set 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 gray line; the ending of zone-specific Bravo shift labels in June 2010 is shown as a vertical dashed gray line. Details of Bravo 1 and Bravo 2 bed locations are in tables B-8.1, B-8.3, and B-8.4.

Alpha for physicians working in Alpha. This percentage is close to 100 percent 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 1 for physicians in Bravo 2 shifts is 0 percent throughout. The percentages for physicians in Bravo 1 shifts are substantially more varied, but the median percentage for physicians in these shifts hovers slightly above 90 percent until January 2010, when plans for the regime change are announced.[37] 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.

VIII. Conclusion

Simply by allowing physicians to choose patients, a self-managed system reduces ED patient lengths of stay by 11–15 percent relative to the nurse-managed system. This effect occurs primarily by reducing a “foot-dragging” moral hazard, in which physicians delay patient discharge to forestall new work. Foot-dragging is sensitive to peer effects, suggesting better information (and social incentives) between peers in the same pod. The self-managed system appears to use this information between peers to improve patient assignment and reduce foot-dragging.

While improving the efficiency of health care delivery through its organization has grown in policy significance (Institute of Medicine 2012), teamwork characterized by mutual monitoring and management is thought to be broadly applicable across many industries (Hackman 1987). A strength of this study is that, via a narrow intervention, it isolates a key feature of teamwork—the joint management of work by workers who have better information—that is likely to be important in a wide variety of settings

[37] Appendix B-8 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, because of 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.

and is distinct from the classic concept of teams as joint production (e.g., Holmstrom 1982).

This study identifies an effect that is likely lower than the potential effect of joint management in other environments. First, its intervention to increase teamwork is remarkably limited. Second, a large amount of information in this ED setting was already publicly available through information technology, and much of the remaining private information in medical care likely remains hidden even to peers in the same location. Third, patient care in the ED is highly tied to decision making by a single generalist physician. Other environments that more heavily feature more flexible division of labor (e.g., matching tasks to skills) could further increase the returns to teamwork.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *J. American Statis. Assoc.* 105 (490): 493–505.

Acemoglu, Daron, Philippe Aghion, Claire Lelarge, John Van Reenen, and Fabrizio Zilibotti. 2007. "Technology, Information, and the Decentralization of the Firm." *Q.J.E.* 122 (November): 1759–99.

Aghion, Philippe, and Jean Tirole. 1997. "Formal and Real Authority in Organizations." *J.P.E.* 105 (1): 1–29.

Alchian, Armen A., and Harold Demsetz. 1972. "Production, Information Costs, and Economic Organization." *A.E.R.* 62 (5): 777–95.

Apker, Julie, Larry A. Mallak, and Scott C. Gibson. 2007. "Communicating in the Gray Zone: Perceptions about Emergency Physician-Hospitalist Handoffs and Patient Safety." *Academic Emergency Medicine* 14 (October): 884–94.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2005. "Social Preferences and the Response to Incentives: Evidence from Personnel Data." *Q.J.E.* 120 (3): 917–62.

———. 2009. "Social Connections and Incentives in the Workplace: Evidence from Personnel Data." *Econometrica* 77 (4): 1047–94.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Q.J.E.* 119 (February): 249–75.

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

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

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. "Does Management Matter? Evidence from India." *Q.J.E.* 128 (February): 1–51.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and Explaining Management Practices across Firms and Countries." *Q.J.E.* 122 (4): 1351–1408.

Caroli, Eve, and John Van Reenen. 2001. "Skill-Biased Organizational Change? Evidence from a Panel of British and French Establishments." *Q.J.E.* 116 (November): 1449–92.

Delarue, Anne, Geert Van Hootegem, Stephen Procter, and Mark Burridge. 2008. "Teamworking and Organizational Performance: A Review of Survey-Based Research." *Internat. J. Management Revs.* 10 (2): 127–48.
Dessein, Wouter. 2002. "Authority and Communication in Organizations." *Rev. Econ. Studies* 69 (October): 811–38.
Donnelly, Laura. 2013. "Minister: NHS Will Collapse If Elderly Bed Blocking Continues." *Telegraph*, May 14.
Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2013. "Truth-Telling by Third-Party Auditors and the Response of Polluting Firms: Experimental Evidence from India." *Q.J.E.* 128 (November): 1499–1545.
Elixhauser, Anne, Claudia Steiner, D. Robert Harris, and Rosanna M. Coffey. 1998. "Comorbidity Measures for Use with Administrative Data." *Medical Care* 36 (January): 8–27.
Fehr, Ernst, and Simon Gachter. 2000. "Cooperation and Punishment in Public Goods Experiments." *A.E.R.* 90 (4): 980–94.
Forster, Alan J., Ian Stiell, George Wells, Alexander J. Lee, and Carl Van Walraven. 2003. "The Effect of Hospital Occupancy on Emergency Department Length of Stay and Patient Disposition." *Academic Emergency Medicine* 10 (2): 127–33.
Fudenberg, Drew, and David K. Levine. 1993. "Self-Confirming Equilibrium." *Econometrica* 61 (May): 523–45.
Gaynor, Martin, James B. Rebitzer, and Lowell J. Taylor. 2004. "Physician Incentives in Health Maintenance Organizations." *J.P.E.* 112 (4): 915–31.
Gibbons, Robert, and Rebecca Henderson. 2012. "What Do Managers Do? Exploring Persistent Performance Differences among Seemingly Similar Enterprises." In *The Handbook of Organizational Economics,* edited by Robert Gibbons and John Roberts, 680–732. Princeton, NJ: Princeton Univ. Press.
Greif, Avner. 1993. "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition." *A.E.R.* 83 (3): 525–48.
Hackman, Richard J. 1986. "The Psychology of Self-Management in Organizations." In *Psychology and Work: Productivity, Change, and Employment,* edited by Michael S. Pallak and Robert O. Perloff. Washington, DC: American Psychological Assoc.
———. 1987. "The Design of Work Teams." In *Handbook of Organizational Behavior,* edited by Jay W. Lorsch, 315–42. Englewood Cliffs, NJ: Prentice-Hall.
Hamilton, Barton H., Jack A. Nickerson, and Hideo Owan. 2003. "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation." *J.P.E.* 111 (June): 465–97.
Holmstrom, Bengt. 1982. "Moral Hazard in Teams." *Bell J. Econ.* 13:324–40.
Ichino, Andrea, and Giovanni Maggi. 2000. "Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm." *Q.J.E.* 115 (August): 1057–90.
Ichniowski, Casey, and Kathryn Shaw. 2003. "Beyond Incentive Pay: Insiders' Estimates of the Value of Complementary Human Resource Management Practices." *J. Econ. Perspectives* 17 (January): 155–80.
Ichniowski, Casey, Kathryn Shaw, and Giovanna Prennushi. 1997. "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines." *A.E.R.* 87 (3): 291–313.
Institute of Medicine. 2012. *Best Care at Lower Cost: The Path to Continuously Learning Health Care in America.* Technical report. Washington, DC: Nat. Acad. Press.
Kandel, Eugene, and Edward P. Lazear. 1992. "Peer Pressure and Partnerships." *J.P.E.* 100:801–17.
Lazear, Edward P. 2000. "Performance Pay and Productivity." *A.E.R.* 90 (December): 1346–61.

Lee, Michael H., Jeremiah D. Schuur, and Brian J. Zink. 2013. "Owning the Cost of Emergency Medicine: Beyond 2%." *Ann. Emergency Medicine* 62 (November): 498–505.e3.

Lerman, Benjamin, and Michael S. Kobernick. 1987. "Return Visits to the Emergency Department." *J. Emergency Medicine* 5 (September): 359–62.

Mas, Alexandre, and Enrico Moretti. 2009. "Peers at Work." *A.E.R.* 99 (1): 112–45.

McHugh, Megan, Kevin Van Dyke, Mark McClelland, and Dina Moss. 2011. "Improving Patient Flow and Reducing Emergency Department Crowding: A Guide for Hospitals." Technical Report no. 11(12)-0094, Agency for Healthcare Research and Quality, Rockville, MD.

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

Milgrom, Paul, and John Roberts. 1988. "An Economic Approach to Influence Activities in Organizations." *American J. Sociology* 94 (suppl.): S154–S179.

Myerson, Roger B., and Mark A. Satterthwaite. 1983. "Efficient Mechanisms for Bilateral Trading." *J. Econ. Theory* 29 (April): 265–81.

Nagin, Daniel S., James B. Rebitzer, Seth Sanders, and Lowell J. Taylor. 2002. "Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment." *A.E.R.* 92 (September): 850–73.

Ostrom, Elinor, James Walker, and Roy Gardner. 1992. "Covenants With and Without a Sword: Self-Governance Is Possible." *American Polit. Sci. Rev.* 86 (June): 404–17.

Proudlove, N. C., K. Gordon, and R. Boaden. 2003. "Can Good Bed Management Solve the Overcrowding in Accident and Emergency Departments?" *Emergency Medicine J.* 20 (March): 149–55.

Rosenbaum, Paul. 2002. "Covariance Adjustment in Randomized Experiments and Observational Studies." *Statis. Sci.* 17 (August): 286–304.

Roy, Donald. 1952. "Quota Restriction and Goldbricking in a Machine Shop." *American J. Sociology* 57 (5): 427–42.

Rubin, Donald B. 1980. Comment on "Randomization Analysis of Experimental Data: The Fisher Randomization Test" by D. Basu. *J. American Statis. Assoc.* 75 (September): 591–93.

Schuur, Jeremiah D., and Arjun K. Venkatesh. 2012. "The Growing Role of Emergency Departments in Hospital Admissions." *New England J. Medicine* 367: 391–93.

Syverson, Chad. 2011. "What Determines Productivity?" *J. Econ. Literature* 49 (June): 326–65.

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

Thompson, David A., Paul R. Yarnold, Diana R. Williams, and Stephen L. Adams. 1996. "Effects of Actual Waiting Time, Perceived Waiting Time, Information Delivery, and Expressive Quality on Patient Satisfaction in the Emergency Department." *Ann. Emergency Medicine* 28 (December): 657–65.

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