Misperception and Technology Adoption

Diego Comin
diego.comin@dartmouth.edu

Jonathan Skinner*
jon.skinner@dartmouth.edu

Douglas Staiger*
douglas.staiger@dartmouth.edu

Department of Economics, Dartmouth College, and NBER
* The Dartmouth Institute for Health Policy & Clinical Practice, Geisel School of Medicine

This draft: 7 November 2018

Abstract:

There are remarkably large differences in the timing of adoption and intensity of use of new technologies. What factors determine such differences? Is it because of higher returns to using the new technology, or systematic misperceptions about their expected returns? We address these questions using a unique registry dataset for a quarter-million patients with implantable cardiac defibrillators (ICDs), a medical device that reduces the risk of sudden cardiac arrest. We develop a structural model of Bayesian learning that allows for misperception of provider skill that can lead to overly optimistic or pessimistic behavior. Briefly, our estimates suggest that for ICDs, the most rapid adopters were overly optimistic, leading both to high utilization rates and for these early innovators, *below-average* returns to the technology. We find that misperception can explain half of hospital-level variation in risk-adjusted mortality and nearly three-quarters of the variation in adoption and use. In addition, the model predicts, correctly, that those hospitals exhibiting the greatest optimism about their own ability are the ones that scale back quickest. These results suggest an important role for misperception (both optimistic and pessimistic) in explaining the wide variation in adoption and use of new technologies, and suggests caution in equating rapid diffusion to productivity gains.

We are grateful for financial support from the National Institute on Aging (P01-AG19783 and U01-AG046830). In addition, we are indebted to Greg Roth, Peter Groeneveld, Kimon Bekelis, and to seminar participants at the World Bank, the University of Chicago, Princeton University, the University of Virginia, The University of Illinois, Emory University, NBER, and the Federal Reserve Bank of Chicago for very helpful comments and suggestions. Weiping Zhou provided essential programming and analysis.

I. Introduction

A central question in the productivity literature is what explains the enormous variation in the adoption and use of new technology (Comin and Hobijn, 2004, 2009; Skinner and Staiger, 2015, Comin and Mestieri, 2018). There is no lack of potential explanations for these variations; Grilliches (1958), for example, emphasized differences in the profitability of hybrid-corn adoption, while Comin and Hobijn (2007) and Caselli and Coleman (2006) rely on heterogeneity across agents in the value of the new technology. Non-adopters may also optimally hold back because they are waiting for the price to decline or are better at the old technology (Jovanovic and Nyarko, 1996), or because they face higher costs from suppliers (Suri, 2011).

A related literature seeking to explain slow diffusion instead as the consequence of poorly informed agents who lack appropriate education or information about potentially profitable innovations (e.g., Foster and Rosenzweig, 1995; Conley and Udry, 2010; Rogers, 2010; Skinner and Staiger, 2007) or time-inconsistency and a lack of commitment devices (Duflo, Kremer, and Robinson, 2008). All of these papers seek to explain why diffusion is so slow despite the clear economic benefits of adopting new technologies, and the implications of this slow diffusion for productivity growth (Comin and Hobijn, 2010, and Comin and Mestieri, 2018).

In this paper, we ask a closely related question: Why are some so quick to adopt a new technology and using it intensively across a wide swath of applications? Nearly all of the previous studies assume that early adopters face greater profitability, better information, and superior relative advantage in the new technology. However, these assumptions have been difficult to test given that there are rarely direct measures of the return for the specific adopter from using the technology, and when they are available, it is difficult to differential between contextual factors (e.g., risk, appropriateness) and others that are intrinsic to the adopter (e.g., skill).

We sidestep these difficulties by using a unique clinical registry for implantable cardioverter defibrillators (ICDs), an expensive medical device that reduces the risk of sudden cardiac arrest, to measure ex post productivity for both early and late adopters in treating patients with congestive heart failure (CHF). This allows us to estimate a Bayesian learning model in which adoption and intensity of use for the technology depends on adopter productivity. We nest within this model the possibility that adopters can exhibit *misperception*, that they could be either overly optimistic or pessimistic about the value of the new technology, or about their own skill in using the new technology. The idea of misperception or more specifically, overconfidence in adoption

has received attention in fields such as psychology (e.g., Moore and Healy, 2008), finance and management (Malmendier and Tate, 2005; Barber and Odean, 2001, Glaser and Weber, 2007; O'Neill, Pouder; and Buchholtz, 1998), industrial organization (Camerer and Lovallo, 1999), and in health care (Berner and Graber, 2008; Cutler et al. 2019), but much less attention in the productivity literature.

Using the Medicare claims data, we find first that that between 2002 and 2005, rates of ICD use nearly doubled nationally, but with considerable variation in the speed of diffusion; in some regions rates quadrupled, while in other regions they barely budged; since that time, however, on average the use of ICDs has scaled back dramatically.[1] As well, we have identified wide variation in productivity across hospitals, with 2-year risk-adjusted mortality exhibiting a standard deviation of 0.031 relative to its mean of 0.218. There has been surprisingly little evidence of "learning by doing," with conditional mortality for ICD procedures declining only slightly, by 0.004, between 2006 and 2013. Finally, we find a positive correlation between the use of ICDs in patients and risk-adjusted mortality, suggesting that patients receiving ICD implants from the most rapid adopters and users of the technology were *more* likely to die following the procedure.

To explain these four empirical facts, we develop an optimizing Bayesian framework where both doctors and patients are heterogeneous and health outcomes are uncertain. Patients differ in the potential benefits from an ICD implant while providers differ in their ability in implanting ICDs. If the providers' perceptions about their skill are unbiased, those with more skill will, on average, adopt earlier and implant more ICDs, while those with lower skill level find fewer patients for which the technology has positive net expected returns (e.g., Currie and MacLeod, 2013; Currie, MacLeod, and Van Parys, 2015; Chandra and Staiger, 2007). Therefore, in the absence of biases in perceived skill, the model predicts that systematic variation in the intensity of use of ICDs is entirely driven by the doctors' skill in applying the technology.

We introduce a new parameter nested in our model that allows for misperception of the doctors' own skill; we deem them to be overly optimistic when the provider gets worse outcomes by going too deeply into the pool of potential patients, and conversely. While our model cannot

[1] The "exnovation" or scaling back of use has been found in other surgical procedures during this period; see Bekelis et al. (2017).

distinguish between the agent's over-optimism about their own skill, or their overly optimistic view of the ICD's value relative to other treatment, we can reasonably rule out an alternative explanation of extrinsic motivation, as in a supplier-induced demand model (Chandra et al., 2011).[2]

Our goal is to fit the theoretical model to the empirical moments of the ICD data. We first calibrate parameters that are common to all hospitals to match aggregate moments related to the cross-sectional distribution of ICD use and mortality rates. Second, given these common parameters, we calibrate the level of true and perceived skills to match observed patterns at the hospital level for ICD use and mortality. We also allow physicians to learn over time about their biases with Bayesian learning about their own skill. A key prediction of the model is that, other things equal, overly optimistic physicians should scale back, with the speed commensurate with the physician's prior distribution of her own skill level. We also estimate the learning model using the hospital-level panel with their associated (initial) misperception parameters, and test the model by predicting mortality and utilization out-of-sample through 2013.

Briefly, we find that misperception is a key driver of ICD use and conditional mortality, explaining half of the variation in risk-adjusted mortality and nearly three-quarters of the variation in rates of adoption and use. Despite the simplicity of the model, it explains roughly 50 percent of the variation in the annual hospital-level change in perceived skill. Finally, the out-of-sample implied ICD use and conditional mortality for 2013 matches the observed variables very closely; the model predicts the drop in ICD use rate from 0.21 per 100 to 0.14, the declines in the ICD standard deviation and cross-sectional correlations between skill and ICD use, and the lack of improvement in aggregate mortality rates.

In sum, these estimates suggest that systematic misperceptions are an important reason why some hospitals experienced such rapid adoption, and others were much slower in adopting. Yet physicians learned from their experience with this new population of patients; those who were the most overconfident also scaled back their use of ICDs most rapidly, leading to a (small) reduction in the conditional mortality rate. Our findings differ from other studies of clinical learning-by-doing (e.g., Jovanovic and Nyarko, 1995; Gong, 2017) in that we find no empirical evidence that

[2] As we discuss in Section 5, in a supplier-induced demand model, some physicians adopt ICDs and over-use them, despite potential harm to their patients, in order to increase their income. In this type of model, there is no learning and scaling back; physicians knew they were harming their patients from the outset, and thus would have had little incentive to scale back.

true skill improved over time. Our results also differ from earlier analyses of diffusion, in which early adopters were the "innovators" and slow diffusers "laggards" (Rogers, 2004). This view has been articulated in health settings by Currie, MacLeod and Van Parys (2015), who find that the most aggressive physicians in treating heart attacks using stents (according to then-current standards) gained the best results. Our finding can be reconciled with theirs by noting that for heart attacks, the new and then unproven technology of stenting turned out *ex post* to have been far more advantageous than expected, while the medical consensus appears to be that ICDs are less successful in practice than first envisioned (McMurray, 2016). That is, heterogeneity in the degree of misperception – whether because of over-optimism or over-pessimism – is likely to be an important explanation for why there is so much variability in the adoption and use of new technologies.

The rest of the paper is organized as follows. Section 2 describes the technology and documents its diffusion patterns, while Section 3 develops the model and Section 4 the analysis. Finally, Section 5 discusses the interpretation of the findings, robustness checks and generalization of our findings to other technologies and sectors in the economy.

2. Implantable Cardioverter Defibrillators (ICDs)

Congestive heart failure (CHF) is a very common illness especially among elderly people (Rogers, 2013), with a prevalence of 5.8 million people in the U.S., far more common than acute myocardial infarctions (or heart attacks), with a prevalence of about 715,000 annually. While heart attacks are sudden medical emergencies treated (often successfully) with a variety of medical interventions, CHF is a chronic illness whose progression can only be slowed by appropriate medical management (Kolata, 2017). The typical progress of CHF is from the New York Heart Association Class I (the least severe) through to Class IV (the most severe), at which point the annual mortality rate ranges between 20-50 percent (Ahmed et al., 2006).

An important risk facing CHF patients is a sudden cardiac arrest, which occurs when the heart suddenly stops functioning, typically because of arrhythmia, or irregular heart rhythm. This causes rapid and unsynchronized heartbeat, leading to little or no blood being pumped from the heart, and a complete absence of a heartbeat (van Reys, 2014). Implantable cardioverter defibrillators (ICDs) are small electronic devices that are surgically implanted in the pectoral region of the chest and connected with wire "leads" to key locations of the heart. These leads

serve two functions. The first is to monitor the rhythm and detect tachycardia (irregular or weak heart beats), and the second is, when necessary, to shock the heart with a strong electrical current, effectively "rebooting" the conduction system. (Popular entertainment shows often show physicians using paddles to administer electrical shocks;[3] ICDs are internal automated versions.) Over time ICDs have become more effective and entailed fewer complications as the size of the ICD shrunk, and the sophistication of the computer programs designed to detect arrhythmias improved.

Initially, ICDs were developed in the 1980s and 1990s for people who had already experienced and survived a cardiac arrest, and were at risk of experiencing another one. As ICDs became more compact and reliable, attention turned to the larger group of people with congestive heart failure (CHF) at risk of cardiac arrest but who had not yet experienced the life-threatening event; for these patients the ICD is deemed "preventive." A large 2005 randomized trial, SCD-HeFT, found substantial mortality benefits of up to 7 percentage point increases in survival 5 years after the procedure (Bardy et al., 2005). It is important to note that ICDs provide no other benefit to patients other than a "reboot" in the case of sudden cardiac arrest; thus mortality as a measure of health outcomes is an apposite measure. Soon after the SCD-HeFT trial, ICDs were allowed by Medicare in the U.S. to be used as a preventive device for patients with weakened hearts (congestive heart failure, or CHF) who had not yet experienced a cardiac arrest, thus expanding dramatically the population of those eligible for ICDs; thus the "adoption" is for the use of an existing technology in a new population, rather than a brand-new technology. We use the Medicare claims data linked to a Centers for Medicare and Medicaid Services (CMS) clinical registry of every ICD implanted during 2006-13 with detailed information on key clinical variables that characterize both appropriateness for treatment, and subsequent risk of mortality.

The SCD-HeFT trial included only the intermediate Class II and Class III CHF patients with low "ejection fractions" or the heart's ability to pump blood to the rest of the body.[4] The reason why the trial was limited to only these two groups was the consensus that for Class I (the least serious) CHF patients, the risks outweighed potential benefits given the rarity of sudden

[3] An example from CSI: New York: https://www.youtube.com/watch?v=_lJCDrYxK9A

[4] As well, the ejection fraction should be 35% or less in patients with Class II or III Heart Failure. Despite the rarity of older patients in the randomized trials, there are no guidelines that recommend against the use of ICDs on the basis of age.

cardiac arrest in this group versus the risks of broken leads or infections, while for the more severe Class IV patients, the heart is so weakened that it can no longer sustain pumping, no matter how many times it reboots. For these patients, ICDs can lead to a series of successive and painful shocks, sometimes delaying an otherwise peaceful demise as the ICD continues to go off until the batteries are drained (Friedrich and Bohm, 2007). Despite these guidelines, a small fraction of ICD procedures were done for those with either Class I or Class IV patients, or for those who had been diagnosed with CHF only recently, and thus have not yet tried medical management. In our analysis, we adjust for these different characteristics, but do not address the more complex problem of whether higher-quality physicians should be more or less likely to follow guidelines.[5]

Finally, to understand the growth and subsequent reduction in the use of ICDs, it is important to rule out the development of a new technology that might have led to a shift away from ICDs. While during this time, there was greater emphasis on adherence to guideline-directed drug prescriptions (e.g., Roth et al., 2016), there was no new innovation or breakthrough developed to reduce mortality among CHF patients (Kolata, 2017).[6]

2.1 Patterns of ICD Diffusion in the Medicare Population

To first study the evolution of ICD use, we use the 100% Medicare claims data for the fee-for-service over-65 Medicare population to derive regional utilization rates that can then be assigned to hospitals (as described below). Because of possible changes over time in coding standards, we develop measures for all new ICD implantations during 2002-13 (thus excluding replacement ICDs because of failed batteries or other reasons), and not simply those designated or coded for preventive purposes.[7] To measure utilization, we use population-based rates at the hospital referral region (HRR) level, of which there are 306 in the U.S.[8] These utilization measures

[5] In the context of our model below, it is possible that higher-skill physicians could still gain good outcomes even for out-of-guideline patients. For this reason we do not include "within guideline" as a quality measure. During the period of analysis, CMS cracked down on hospitals billing for out-of-guideline patients; however, these changes had no impact on risk-adjusted mortality, which adjusts for all characteristics of the patient.

[6] Recently a new treatment was developed for mitral valve regurgitation; see Kolata (2018).

[7] We begin the analysis using the claims data in 2002, when the sample of Part B claims data relevant for analysis is 20% of all fee-for-service enrollees; the sample rises to 40% in 2003-05, and becomes 100% thereafter. We use CPT 33249 rather than in-hospital DRG codes to measure incidence.

[8] HRRs were first developed by the Dartmouth Atlas project in the 1990s to create regions based on the migration patterns of individuals to their hospitals. Thus HRR boundaries will often follow (e.g.) interstate highways and

are based on the residence of the patient; if a resident of the Memphis HRR received their ICD in Atlanta, the ICD would be assigned to the Memphis HRR rather than to Atlanta's.

Regions may differ in their use of ICDs because of a greater prevalence of disease. For this reason, we adjusted all HRR-level rates, in each year, using a linear-probability year-specific risk adjustment model.[9] We include as predictors at the individual level five-year age brackets (and a category of 85+), sex, race (black, white, and other), at least one physician visit with a diagnosis of congestive heart failure, dual eligibility with Medicaid (an individual indicator of serious illness, poverty, or both). At the ZIP code level we included poverty rates and income (from the 2010 Census) and at the county level smoking, obesity, and diabetes based on Behavioral Risk Factor Surveillance System (BRFSS) data; these latter health behavior measures are highly predictive of regional mortality rates (Wennberg et al., 2014). The regression estimates, presented in Table A.4 for three selected years (2002, 2006, and 2013), indicate that individual attributes are important risk adjusters – particularly the diagnosis of CHF – but that the measures of health behaviors are less important.[10]

In Figure 1, we present risk-adjusted population-based rates of ICD use by HRR between 2002-13 for the U.S., and for selected regions, with an emphasis on the regions adopting most rapidly. Note that between 2002 and 2005, average ICD use increased from 0.12 per 100 Medicare enrollees to 0.23, a near doubling of average rates, with a decline in rates to 0.15 per 1000 by 2013.

Some part of this increase could have been because of a "stock-flow" issue; the stock of patients newly eligible for the ICD could have lead to an uptick in utilization for 2006, generating the increased rates. As we discuss in more detail below, the ICD registry data includes the duration of the CHF, so we might expect that the duration of CHF for patients getting an ICD in 2006, for example, would be longer than for patients in 2013. This hypothesis would imply that a typical ICD patient would have experienced CHF for a longer period of time in 2006 compared to 2013.

---

cross state lines. Each HRR includes a major tertiary hospital that performs neurosurgery and cardiac surgery. We use HRRs rather than the smaller hospital service areas (HSAs) for better sample precision.

[9] We use year-specific risk adjustment regressions because average rates of use vary so much by year. Alternatively, we could have specified a logistics or probit over all years, but the computational requirements (e.g, estimating 3500 individual coefficients in a sample of nearly ½ billion observations) would have been excessive.

[10] We are probably over-adjusting because more aggressive physicians are likely to both be more likely to diagnose CHF for "gray area" patients, and prescribe ICDs.

However, if anything the opposite is found; those with CHF duration more than 9 months actually rose during the period, from 72 to 82 percent.

As suggested by Figure 1, there is widespread variability in rates of utilization. Three of the most rapid adopters were Munster IN (from 0.14 in 2003[11] to 0.37 in 2006), Mason City IA (from 0.13 in 2004 to 0.49 in 2007), and Terre Haute IN (from 0.12 in 2003 to 0.55 in 2005). By contrast, many larger metropolitan regions exhibited much smaller increases, with rates remaining low (e.g., Seattle, Manhattan, as well as other cities such as Los Angeles not reported in the Figure) throughout the entire period.

Figure 2 provides a map for the entire U.S. of 2006 ICD utilization rates by HRR. This figure confirms the geographic disparity in the use of ICDs across the entire U.S., with a 10-fold difference between Victoria, Texas (0.03 in 2006) to Terre Haute (0.50 in 2006).[12] Finally, we explore whether the decline in ICD use is driven by those hospitals that adopted ICDs more intensively or those that adopted them less. Figure 3 plots the change in hospital-level ICD use rate between 2002 and 2005 (x-axis), and between 2006 and 2013 (y-axis). The cross-hospital correlation between initial and subsequent change in ICD rates is strongly negative (-.40, $p < .001$), showing that the hospitals with the most rapid initial growth also experienced the most rapid decline. Despite the overall decline in the ICD use rate, however, the coefficient of variation (the standard deviation divided by the mean) declined only slightly, from 0.31 in 2006 to 0.28 in 2013.

While population-based rates of ICD utilization are drawn from HRRs, we seek to estimate our model at the level of the hospital that performs the ICD.[13] We do this by assigning to each patient their HRR-level utilization measure (as described above). For example, if a hospital in the Boston area draws from the Boston, Providence, and Portland ME HRRs for their ICD patients, the hospital-specific rate of ICD utilization will be a weighted average of those three HRR rates; this is shown in a schematic in Figure 4a.

[11] The corresponding rate in 2002 is suppressed because the numerator comprises fewer than 11 observations, the CMS limit for reporting data.

[12] One might be concerned with small-sample bias in these relatively small HRRs, but the patterns show a strong temporal trend; high rates in 2006 are matched (or even exceeded) by high rates in 2005 and 2007.

[13] Measures of ICD intensity in utilization requires both a numerator (the number of ICDs implanted in a given year) and a denominator (the number of potential patients). While regions are well suited to calculate both numerator and denominator (e.g., as done in the HRR-level analysis above), calculating the denominator of a given hospital, particularly in a city with multiple hospitals, is exceedingly difficult.

2.2. Variation in Health Outcome Following ICD Implantation

A rare luxury in studies of technology diffusion is to have access to information on the performance of adopters after adopting the technology. When CMS approved the use of ICDs for preventive purposes, it was done with the understanding that hospitals would send detailed clinical information about the patient to CMS. We use this 100% registry, linked to the Medicare denominator file for people age 65+, during 2006-13, which allows us to calculate mortality rates based on Medicare denominator files available through 2015.[14] The registry includes detailed information on the registry that includes whether the ICD was for patients with CHF (e.g., preventive), their risk class (I through IV) as well as ejection fraction and many other clinically relevant factors such as having ventricular tachycardia, family history of cardiac arrest, the exact ejection fraction, and other measures, along with the identity of the hospital performing the procedure.[15] These data are far more detailed than what could ever be recovered from Medicare billing claims. To estimate outcomes, we focus on a relatively homogenous group of CHF patients who have never had an ICD implanted; we implicitly assume that the hospital-specific mortality effect estimated using these patients is similar to the effect for other patients receiving an ICD.[16]

Ideally, we would like to measure true treatment effects; the benefit of an ICD relative to the status quo of medical management for CHF. However, because we do not observe patients who did not receive and ICD in our registry data, our estimates and modeling are specific to mortality rates only among those treated; we discuss this concern, and how we address it, in the modeling section in Section 4.

Table 1 provides summary statistics of the ICD sample (N = 253,613). The average age among the Medicare enrollees (all of whom are 65+) is 74.5, and just 28 percent are female. Note

[14] Mortality data are drawn from the 2006 mortality data in www.dartmouthdiffusion.org.

[15] One complexity associated with identifying hospitals is that in some cases, the hospital was not identified; only the NPI for the provider who performed the procedure. We are grateful to Andrea Austin for providing a cross-walk from ICD-capable providers to the hospital where they performed the plurality of procedures, which we used to create our dataset.

[16] We recognize that some hospitals may include more than one cardiologist or electrophysiologist who performs the procedures, but identifying pure physician effects from the hospital-level team that both implants and also maintains the ICD is problematic.

that the mortality rate barely budged between 2006 and 2013. We also include summary statistics for additional covariates from the registry, including the ejection fraction, prior cardiac arrest, family history, prior heart attack, and other variables.

Hospital-level risk-adjusted mortality is modeled using the following hierarchical structure:

$$M_{jit} = \Psi_{it} + X_{jit}\beta + \zeta_{jit} \quad (1)$$

$$\text{where } \Psi_{it} = ICD\_rate_{it}\Gamma + \theta_i + v_{it} \quad (2)$$

The first equation is at the patient level, where mortality (M_{jit}) for patient j treated at hospital i in year t is a binary variable that depends on characteristic of the patient (X_{jit}), and a hospital effect (Ψ_{it}). At the hospital level, the hospital effect in turn depends on the hospital-level utilization rate of ICDs in that year (ICD_rate_{it}) plus a random hospital effect (θ_i) and a random hospital-year effect (v_{it}). We allow the hospital effect of mortality to depend on the utilization rate of ICDs because that is implied by our model. We are particularly interested in the variance of Ψ_{it} and its covariance with the hospital's ICD utilization, which depends both on the predictable characteristics of the hospital, Var($ICD\_rate_{it}\Gamma$), as well as the provider-specific error term Var(θ_i). Our preferred specification is a hierarchical random-effects model, which provides estimates of the key parameters ($\Gamma, Var(\theta_i)$) and also estimates of the individual hospital effects, $\widehat{\Psi}_{it} = ICD\_rate_{it}\hat{\Gamma} + \hat{\theta}_i + \hat{v}_{it}$, where we use best linear unbiased predictions for the hospital and hospital-year random effects (e.g., "shrink" the estimate of the provider residual towards the fitted value $ICD\_rate_{it}\hat{\Gamma}$ depending on the sample size of the provider). We focus on random-effects models, but in sensitivity analyses we also consider least-squares regressions and models with provider-level fixed effects. Because we wish to estimate the hospital-specific effect on mortality of ICD relative to medical management, in some specifications we add hospital-level controls (to X) to proxy for quality of medical management such as patient volume and the use of guideline-consistent medical treatment for CHF patients.,

The benefits inherent in ICD implantation arise only after several years (Bardy et al., 2005) so we focus on both 1-year and 2-year mortality. For the random-effects model, we estimate the distribution of Ψ_{it} in Figure 4, which shows the risk-adjusted one-year variation in hospital-specific mortality rates. As is clear from Figure 4, there is significant variation in conditional

mortality rates across hospitals; the standard deviation of one-year conditional mortality rate across hospitals is 2.2 percentage points, with mortality rates in high-mortality hospitals that are twice as high as those in low-mortality ones.[17] Finally, Figure 6 presents the evolution of one-year conditional mortality over time for the U.S. Conditional one-year mortality declined slightly from 12.8% in 2006 to 12.0% in 2013.

2.3 The Correlation Between ICD Diffusion and Mortality

A natural question that should help us understand the drivers of ICD diffusion is the relationship between ICD use and conditional mortality. To this end, we combine the utilization data (Section 2.1) and the outcome estimates (Section 2.2) to compute the correlation between conditional mortality and ICD utilization. Rather than report mortality rates at the hospital level, we instead convert them back to the HRR level based on the residence of the patient, as shown schematically in Figure 4b. Figure 7 shows the correlation between the average (2006-13) ICD utilization rate, and the fully risk-adjusted relevant hospital-level 2-year mortality. The correlation coefficient is 0.15 for one-year mortality, and 0.11 for two-year mortality. The graph also identifies several of the more interesting regions; in particular those regions exhibiting both low mortality rates and low use of ICDs (the Minneapolis-St. Paul HRR); while others exhibit high rates of ICD use, coupled with high rates of mortality, such as Miami, Terra Haute IN, and Munster, IN. That Munster is an outlier may be explained in part by a specific cardiologist who was sued by for inappropriate cardiac surgery and ICD placement (Creswell, 2015).

In Table 2, we report summary estimates of the OLS, random effect, and fixed-effect models, limited to just two-year mortality; regression results are reported in the Appendix for OLS in Table A.1, random effects in Table A.2, and hospital fixed-effects in Table A.3 that also include one-year mortality. As shown in Table 2, there is a consistent positive correlation and significant correlation between the rate of use of ICDs in a given year, and risk-adjusted mortality rates, in both the OLS and random-effects model, suggesting in the reduced form that patients of the most rapid diffusers experience worse outcomes. The point estimates are much smaller and not significant in the fixed-effect model; this is because most of the identification is from cross-sectional variation.

[17] Recall that these estimates are derived from the random-effects model, and are therefore already shrunken towards the mean; a fixed-effects model would have exhibited even more variability.

As sensitivity analysis, we also include the log of annual volume of all ICD performed at the hospital for the over-65 population (including non-CHF patients), to adjust for the conventional finding that higher-volume hospitals yield better outcomes. The coefficients on these variables are as expected; an increase in log-volume of 1 leads to a 1.4 percentage-point decline in 2-year mortality in the random-effects model (Column 4 of Table 2). Another sensitivity analysis included quality measures for regional medical management specific to ICD patients (Roth et al., 2016). The idea is that for hospitals with poor ICD outcomes, ICDs may provide a better option than the alternative – medical management – if physicians at that hospital do an even worse job of medical management (e.g., Chandra and Staiger, 2007). While the coefficient estimates are in the expected direction (worse medical management and low volume are both associated with worse outcomes for ICD patients), including them as controls does not significantly attenuate the coefficient on utilization.

To sum up, we find wide variation in rates of diffusion across the U.S. with regard to ICD use; reversion to the mean with regard to utilization, in the sense that regions with the most rapid growth were most likely to "exnovate" or scale back on their use (Bekelis et al., 2017); wide variability in ICD mortality rates across hospitals, and a positive correlation between utilization and mortality. We turn next to developing a model that can potentially explain these empirical patterns.

3. The Model

Our goal is to develop a model of technology adoption/use that helps us understand the empirical patterns of ICD utilization and conditional mortality. It builds on an optimizing Bayesian framework where both physicians and patients are heterogeneous and health outcomes are uncertain. Patients differ in the potential benefits from an ICD implant while physicians and their teams differ in their ability in implanting ICDs. Additionally, we recognize that physicians may have biased perceptions about their true ability and about the intrinsic value of ICD implants.

We use the model in three different ways. First, it helps us study how the decision to implant an ICD and the mortality conditional on an ICD implant depend on the health provider's perceived and actual skills. Second, it helps us rationalize the risk adjustments made in the empirical measures of ICD use and mortality conditional on ICD implant. By formalizing this link we can better ascertain the required assumptions for our measures to be unbiased. Finally, we

extend the model by allowing physicians to learn from patient outcomes. This allows us to study the role of learning in the evolution of ICD use and mortality by comparing the out-of-sample model predictions with the actual data.

We note that while our model is couched in terms of physician decisions, our data is at the level of the hospital. This is because ICD procedures are typically team efforts; nurses, electrophysiologists, contribute at various stages to better or worse outcomes. For many hospitals, there is only one primary ICD-capable physician, in which case this assumption is innocuous; for larger hospitals we will be blending the choices of two or more physicians.[18]

3.1 Static setting

We begin with the decision problem from the perspective of the physician. There is a continuum of patient types j that differ in their potential value of the ICD implant and of the alternative treatments. The value of the implant for patient j depends on the patient type v_{Xj} and on the doctor's skill level a_i; given that the only goal of the ICD is to keep the patient alive in the event of sudden cardiac arrest, v_{Xj} can reasonably be viewed as survival. In particular, the value for a patient after an ICD implant is[19]

$$v_{Xj} + a_i \tag{3}$$

We, as econometricians, do not observe the patient type but only some patient characteristics X_j, which we assume to be a zero-mean vector in the population; X_j is also observable by the doctor. In addition, the patient type is also defined by a component, v_j, that is unobservable to us.

$$v_{Xj} = X_j * \gamma + v_j \tag{4}$$

The value of patient j if she receives an alternative treatment, w_{Xj}, also has a component that depends on the patient observable characteristics, $X_j * \delta$, and a component that is unobservable to us, w_j:

$$w_{Xj} = X_j * \delta + w_j \tag{5}$$

[18] We also assume that the ICD-capable physician makes the final decision about which patients to choose. The networks of primary care physicians and how they "feed" patients to the ICD-capable hospitals may also affect choices of patients; see for example Moen et al. (2018).

[19] Without loss of generality, we normalize the costs of implanting an ICD to 0.

Let μ_{Xj} denote the difference between the patient's potential value from receiving an ICD implant v_{Xj} and her value from alternative treatments, w_{Xj}. That is

$$\mu_{Xj} \equiv v_{Xj} - w_{Xj} = X_j * \beta + \mu_j \quad (6)$$

The distribution of patient's net value from treatment conditional on her observable characteristics is normal. In particular,

$$\mu_{Xj}|X_j \sim N(X_j\beta + \bar{\mu}, \sigma_\mu^2), \quad (7)$$

where $\bar{\mu}$ is the population mean of μ_j. The precision of the prior of μ_j is denoted by $\rho_\mu = \frac{1}{\sigma_\mu^2}$. The two components of the net value from treatment, v_j and w_j, are normally distributed and, for the time being, we assume that they are independent; we relax this assumption below. Therefore, $\sigma_\mu^2 = \sigma_v^2 + \sigma_w^2$.

Doctor's information structure and priors. We make three assumptions about how doctors perceive the patient's type, μ_{Xj}, and their skill level, a_i. First, doctors do not directly observe μ_{Xj}. They just observe an imperfect signal of s_{Xj} that takes the form:

$$s_{Xj} = \mu_{Xj} + \varepsilon \quad (8)$$

where ε is normal with mean 0 and variance σ_ε^2.

Second, we allow doctors to have a biased prior on the net value of ICD implants in population. In particular, the prior distribution of $\mu_{X_j}^i|X_j$ for doctor *i* is

$$\mu_{X_j}^i|X_j \sim N(X_j\beta + \bar{\mu}_i, \sigma_\mu^2), \quad (9)$$

where $\bar{\mu}_i - \bar{\mu}$ is the bias in doctor *i*'s perception of the average net value of ICD implants. If $\bar{\mu}_i - \bar{\mu}$>0, doctor *i* believes that ICD implants are on average better than what they actually are.

Third, doctors do not know their true skill, a_i; $\tilde{a}_i^p$ denotes the mean of the doctor's prior distribution of a_i. We refer to $\tilde{a}_i^p$ as perceived skill. The gap between the perceived and true skill is the misperception bias, $\tilde{o}_i$. If $\tilde{a}_i^p > a_i$ the physician is overly optimistic (or overconfident) with regard to her skill, while if $\tilde{a}_i^p < a_i$ she is under-confident or pessimistic. If $\tilde{a}_i^p = a_i$ the doctor is unbiased.

Treatment decision. Doctor *i* will implant an ICD to patient *j* if the expected value from implanting an ICD given the observable information X_j and his private signal s_{Xj} is greater than the expected value from alternative treatment. That is,

$$E_i\left[v_{Xj} - w_{Xj} + a_i|s_{Xj}, X_j\right] \geq 0. \quad (10)$$

Given the information structure, the posterior distribution of the patient's net type is

$$\mu^{i}_{X_j}|s_{Xj}, X_j \sim N(\bar{\mu}^{p}_{i,j}, \frac{\sigma^2_{\mu}\sigma^2_{\varepsilon}}{\sigma^2_{\mu}+\sigma^2_{\varepsilon}}) \tag{11}$$

where the posterior mean is

$$\bar{\mu}^{p}_{i,j} = X_j\beta + (1-\alpha)\bar{\mu}_i + \alpha s_j, \tag{12}$$

with $\alpha = \frac{\sigma^2_{\mu}}{\sigma^2_{\mu}+\sigma^2_{\varepsilon}}$, and with $s_j = s_{Xj} - X_j\beta$ being the signal net of the observable characteristics. Using the expression (11) for the posterior mean of $\mu^{i}_{X_j}|s_{Xj}, X_j$, condition (10) becomes

$$X_j\beta + (1-\alpha)\bar{\mu}_i + \alpha s_j + \tilde{a}^{p}_{i} \geq 0, \tag{13}$$

which using the definitions of $\bar{\mu}_i$ and a^{p}_{i}, implies

$$\overbrace{X_j\beta + (1-\alpha)\bar{\mu} + \alpha s_j + a_i}^{Unbiased\ net\ benefit} + \overbrace{(1-\alpha)(\bar{\mu}_i - \bar{\mu}) + \tilde{o}_i}^{Physician\ bias} \geq 0 \tag{14}$$

Expression (14) decomposes the net benefits from implanting an ICD perceived by the physician into two components. The first component -- labelled unbiased net benefit – is the net benefit that would perceive a physician that does not misperceive her skill or the net value of ICD implants. The second term captures the physician's misperceptions about the improvement resulting from the ICD. Note from (14) that from the perspective of the decision to implant an ICD, these doctor biases are isomorphic. Therefore, we can subsume the doctor biases into a unique term that we denote by b_i. Using this notation, we can rewrite the condition for an ICD implant as

$$X_j\beta + (1-\alpha)\bar{\mu} + \alpha s_j + a_i + b_i \geq 0 \tag{15}$$

Expression (15) implies that doctors implant an ICD, if they receive a signal s_j greater than a threshold $s(a^{p}_{i}, X_j)$ defined by:

$$s_j \geq s\left(a^{p}_{i}, X_j\right) \equiv -\frac{(1-\alpha)}{\alpha}\bar{\mu} - \frac{X_j\beta}{\alpha} - \frac{a^{p}_{i}}{\alpha} \tag{16}$$

where the variable $a^{p}_{i} = a_i + b_i$ is the sum of the doctor's true skill plus her biases due to her misperceptions in her true ability and on the net value of ICDs. For brevity, we refer to a^{p}_{i} as perceived skill but the reader should remember that this variable also includes the bias in the doctor's prior about $\bar{\mu}$.

ICD usage. The probability of implanting an ICD for a doctor with perceived skill a^{p}_{i} in a patient with observable characteristics X_j is

$$\Pr\left(\ ICD = 1 | a^{p}_{i}, X_j\right) = \int_{s(a^{p}_{i}, X_j)}^{\infty} f(s)\, ds, \tag{17}$$

where $s\left(a_i^p, X_j\right)$ is defined by equation (4) and where *f(.)* is the pdf of the signal s_j. That is, it is a normal distribution with mean $\bar{\mu}$, and variance $\sigma_\mu^2 + \sigma_\varepsilon^2$.

Proposition 1 (Determinants of diffusion). Ceteris paribus, the use of ICDs increases with perceived skilled, a_i^p.

Proof:

$\frac{\partial \Pr(\, ICD=1|a_i^p)}{\partial a_i^p} = - f\left(s\left(a_i^p, X_j\right)\right) * \frac{\partial s(.)}{\partial a_i^p} > 0$, because, from expression (4), $\frac{\partial s(.)}{\partial a_i^p} < 0$.□

Intuitively, the threshold signal required to implant an ICD decreases with perceived skill. Therefore, doctors with a high perceived skill are more likely to observe a patient' signal above their threshold. Note that what matters for the incidence of ICDs is the doctors perceived signal, a_i^p. (Recall that $a_i^p = a_i + b_i$.) Therefore, for a given skill, a_i, the use of ICDs increases with greater optimism (or overconfidence), b_i. Similarly, for a given level of b_i, higher (true) skill induces a greater use of ICDs. Note also that expression (17) shows that the only doctor-specific parameter that affects the ICD use rate is the perceived skilled of the doctor, a_i^p. Therefore, given the population parameters that define the distribution of signals, $\bar{\mu}$ and α, we could (and will) use the observed ICD use rates to infer the doctor's perceived skill level.

Outcomes. In our dataset we have information on the mortality rate conditional on an ICD implant. To use this information, we need to translate what death means in our model. Naturally, the event of death (in the near term) should be associated with a low ex-post value for the patient. It also seems reasonable that a death that occurs further in the future is associated with a higher ex-post utility for the patient. By applying this logic, we can establish a mapping between mortality and utility. In particular, we interpret the death of the patient within *x* years as an ex-post utility below a threshold $\underline{\kappa}_x$, where $\underline{\kappa}_x$ is increasing in *x*.

The *x*-years mortality rate conditional on an ICD implant *for a doctor* with perceived skill, a_i^p, and actual skill, a_i, is:

$$Pr\left(v_j + a_i \leq \underline{\kappa}_x \middle| ICD = 1, a_i^p, a_i, X_j\right) = \frac{\Pr\left(v_j \leq \underline{\kappa}_x - a_i \cap\ ICD = 1\right)}{\Pr\left(\, ICD = 1|a_i^p, X_j\right)} =$$

$$\frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_\varepsilon(\varepsilon\prime) f_\omega(\omega\prime)\left(\int_{s\left(a_i^p, X_j\right)+\omega\prime-\varepsilon\prime}^{\underline{\kappa}_x - a_i} f_v(v\prime)dv\prime\right)d\omega\prime d\varepsilon\prime}{\int_{s\left(a_i^p, X_j\right)}^{\infty} f_s(s\prime)ds\prime} \qquad (13)$$

where $f_\varepsilon(.)$ is the pdf for ε, $f_v(.)$ and $f_\omega(.)$ are the pdf for patient's type v_j and ω_j, and $f_s(.)$ is the pdf for the signal *s*. While utilization is affected only by perceived skill, conditional mortality is affected by both the doctor' true skill and perceived skill. This observation is the basis to identifying the true skill levels.

Proposition 2 (Determinants of mortality conditional on ICD implant). (i) The probability of death conditional on implanting an ICD increases with the doctor's misperception, b_i, and (ii) Skill has an ambiguous effect on the conditional mortality rate.

Note that in both of these results we do not condition on patient's characteristics, other than for the fact that they have received an ICD.

Proof: The proofs are as follows:

(i)

$$\frac{\partial \Pr(v_j + a_i \le \underline{\kappa}_x | \, ICD = 1, a_i^p, a_i)}{\partial b_i} =$$

$$\left[1 - Pr(v_j \le \underline{\kappa}_x | \, ICD = 1, a_i^p, a_i)\right]\left(-\frac{\partial \bar{s}}{\partial o_i}\right)\left[\frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_\varepsilon(\varepsilon\prime) f_\omega(\omega\prime) f_v(s(a_i^p, X_j) - \varepsilon\prime) d\omega\prime d\varepsilon\prime}{\int_{-\infty}^{\infty} f_\varepsilon(\varepsilon\prime)\int_{-\infty}^{\infty} f_\omega(\omega\prime)\left(\int_{s(a_i^p, X_j)+\omega\prime-\varepsilon\prime}^{\infty} f_v(v\prime) dv\prime\right) d\omega\prime d\varepsilon\prime}\right] > 0$$

(14)

Both the first and third terms are positive, but the key is the middle expression; that when overconfidence rises, the "hurdle" point at which the physician does the procedure declines, thus expanding the number of patients for which the net benefit is negative.

(ii)

$$\frac{\partial \Pr(v_j + a_i \le \underline{\kappa}_x | \, ICD = 1, a_i^p, a_i)}{\partial a_i} = -\left[\frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_\varepsilon(\varepsilon') f_\omega(\omega') f_v(\underline{\kappa}_x - a_i) d\omega' d\varepsilon'}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_\varepsilon(\varepsilon') f_\omega(\omega')\left(\int_{s(a_i^p, X_j)-\varepsilon}^{\infty} f_v(v') d\, v'\right) d\omega' d\varepsilon'}\right]$$

$$+ \frac{\left[1 - Pr(v_j \le \underline{\kappa}_x | \, ICD = 1, a_i^p, a_i)\right]}{\alpha}\left[\frac{\int_{-\infty}^{\infty} f_\varepsilon(\varepsilon\prime) f_\omega(\omega\prime) f_v(s(a_i^p, X_j) - \varepsilon\prime) d\varepsilon\prime}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_\varepsilon(\varepsilon\prime) f_\omega(\omega\prime)\left(\int_{s(a_i^p, X_j)-\varepsilon}^{\infty} f_v(v\prime) dv\prime\right) d\omega\prime d\varepsilon\prime}\right] \qquad (15)$$

Expression (8) shows that skill affects mortality by improving the outcomes for patients who would have been treated anyway (first term), but also by bringing in more patients with net

benefit, but whose underlying mortality probability could be higher as well (second term).[20] As a result, the net effect of skill on conditional mortality is ambiguous. Of course, if we controlled for patient characteristics, expression (8) collapses to its first term and skill exerts an unambiguously positive effect on health outcomes.

3.2 Dynamics through learning

We explore the dynamic properties of the model by allowing doctors to learn about their true skill. The learning problem we pose is one where doctors are uncertain both about the level of their skill but are certain about the precision of the signals they receive. Additionally, doctors may differ in the precision of their priors about their skill.

We start by describing the nature of the signals and the priors. After implanting *n* ICDs for patients newly covered by the CMS rules, doctors receive *n* imperfect signal, $\{s_{ik}^a\}_{k=1}^n$. Signals are random draws from a normal distribution with unknown value of the mean a_i and known value of the precision ρ_{sa}. In particular, the signal $s_{ik}^a = a_i + \xi_{ik}$, where the noise term ξ_{ik} is distributed according to a normal with zero mean and precision ρ_ξ. The doctor's prior of the distribution of a_i is normal with mean a_i^p and precision τ_i such that $\tau_i>0$ and $-\infty < a_i^p < \infty$.

Note that, in addition to the bias in the gap between true and perceived skill (overconfidence), doctors may have a different bias now in the precision of the conditional prior distribution of skill. For any given level of perceived skill, some doctors may be very confident about the accuracy of their estimate (stubborn) while others may be too unsure (insecure). One way to measure this bias in confidence is by the difference between the precision in the prior distribution of skill, τ_i, and the precision of skill implied by the signal s_{ik}^a (that is, the signal net of noise), $\rho_a = \frac{\rho_{sa}\rho_\xi}{\rho_\xi - \rho_{sa}}$. Therefore, if $\tau_i>\rho_a$ a physician is stubborn and if $\tau_i<\rho_a$ she is insecure.

The following Lemma, helps us characterize the evolution of perceived skill over time.

Lemma 1 (Posterior distribution of skill) The posterior distribution of a_i is normal with mean $a_i^{p\prime}$ and precision $\tau_i+n\rho_{sa}$, where

$$a_i^{p\prime} = \frac{\tau_i a_i^p + n\rho_{sa}\overline{s_i^a}}{\tau_i + n\rho_{sa}} \tag{16}$$

[20] This is one reason why some highly-skilled physicians may appear to be lower quality; because they end up with the most difficult patients.

and $\overline{s}_i^a = \dfrac{\sum_{k=1}^{n} s_{ik}^a}{n}$

Proof: See De Groot (1971), page 167. □

Lemma 1 describes the evolution in perceived skill. Subtracting a_i in both sides of expression (9) and substituting the tildes by time subscripts we obtain

$$a_{i,t+1}^p - a_{i,t}^p = -\alpha_{it}^a * \left(a_{i,t}^p - a_i\right) + \alpha_{it}^a * \sum_{k=1}^{n} \xi_{it} \qquad (17)$$

where

$$a_{it}^a = \frac{n\rho_{sa}}{\tau_i + n\rho_{sa}} > 0 \qquad (17)$$

is the learning coefficient.

Replacing in $\bar{a}_{i,t}^p - a_i$ by $o_{i,t}$, we obtain[21]

$$a_{i,t+1}^p - a_{i,t}^p = -\alpha_{it}^a * b_{i,t} + \alpha_{it}^a * \sum_{k=1}^{n} \xi_{it} \qquad (18)$$

Because α_{at} is, in principle, independent of the bias, equation (12) shows that the larger the bias in perceived skill, the larger the expected correction in the perceived skilled, and conversely.

Adding and subtracting a_i to the left-hand-side of (12), we can express the law of motion for misperception as

$$b_{i,t+1} - b_{i,t} = -\alpha_{it}^a * b_{i,t} + \alpha_{it}^a * \sum_{k=1}^{n} \xi_{it} \qquad (19)$$

Equation (19) shows that learning induces mean-reversion in the level of doctor misperception, whether overconfidence or under-confidence.

The speed of learning in our model is captured by the coefficient α_{it}^a. Expression (17) shows that α_{it}^a varies across doctors. α_{it}^a decreases in the precision of the prior precision of skill (τ_i), and increases in the precision of the signal (ρ_{sa}) and in the number of signals/ICD implants

[21] Subtracting and adding a_i to the right-hand side of (14) we obtain the following expression for the evolution of overconfidence:

$$\Delta b_{i,t+1} = -\alpha_{at} * b_{i,t} + \alpha_{at} * \xi_t \text{ where } \Delta b = b_{i,t+1} - b_{i,t}.$$

(n).[22] Other things equal, a more stubborn doctor will learn more slowly (i.e. will have a lower α_{it}^{a}).

4. Analysis and Model Estimation

We next use the model to study the determinants of empirical patterns of ICD use and conditional mortality documented in Section 2. Our strategy has 3 steps. First, we use the values of aggregate moments to calibrate parameters of the distributions of patient type, and doctor true and perceived skill. These parameters are common across hospitals and impact the decision rules of doctors and the outcomes from implanting ICDs. Second, using the common parameters and the hospital-level data on usage rate of ICDs and mortality conditional on ICD, we identify the hospital-level true and perceived skill, a_{it} and a_{it}^{p}. Third, we use the identified parameters to: (i) explore determinants of ICD use and conditional mortality in the cross-section and time series; (ii) conduct counterfactual exercises of "turning off" misperception, and (iii) using out-of-sample approaches, study the ability of the learning model to predict the evolution of misperception, ICD use and mortality over time.

4.1 Identification

Unit of observation. As noted earlier, while our model captures decision-making process of individual physicians, our unit of observation is the hospital, where treatment decisions and quality are typically determined by small teams, often lead by a cardiologist or electrophysiologist.

Aggregate parameters. We start by calibrating the parameters that are common across hospitals. Without loss of generality can normalize the average skill in population, $\bar{\boldsymbol{a}}_{\boldsymbol{i}}$, and the average utility of a patient with heart failure in the absence of ICD treatments, $\bar{\boldsymbol{w}}$, to 0. These

[22] Furthermore, α_{it}^{a} evolves over time. In the special case where $n\rho_{sa}$ is constant in any given hospital, we can use Lemma 1 to derive the following difference equation for α_{at}:

$$\alpha_{at+1} = \frac{\alpha_{at}}{1+\alpha_{at}}$$

The solution to this difference equation is

$$\alpha_{at} = \frac{\alpha_{0i}}{1+t*\alpha_{0i}}$$

where $\alpha_{0i} = \frac{\sigma^2_{a_0^p}}{\sigma^2_{a_0^p}+\sigma^2_{i\xi}}$ is the learning coefficient in the initial period.

parameters are isomorphic to $\bar{v}$ in the ICD use equation (5), and to $\underline{\kappa_x}$ in the mortality rate equation (6).

An important aspect of the calibration is to bridge the conceptual gap between the units in the model (i.e., utility) and in the outcomes we observe (i.e., mortality after the ICD implant). We do this by calibrating the thresholds $\underline{\kappa_x}$ to match the unconditional mortality rates for patients with congestive heart failure (CHF).[23] Specifically, we set $\underline{\kappa_x}$ so that the cdf of w_j is equal to the *x*-years mortality of patients with CHF.

These leaves 7 parameters to calibrate, the average level in population of misperception and the value of ICDs ($\bar{b}$ and $\bar{v}$), and the variance in population of three patient-level parameters, v_j, w_j, ε_j, and two hospital/doctor level parameters a_i and b_i ($\sigma_v^2, \sigma_w^2, \sigma_\varepsilon^2, \sigma_a^2, \sigma_b^2$). To calibrate these parameters, we use 8 moments: the mean and variance ICD use rate across hospitals, the mean and variance conditional 1- and 2-year mortality across hospitals, and the cross-hospital correlations between the ICD use rate and the 1- and 2-year conditional mortality rates.[24] Note that our system is over-identified.

A narrative for the model identification is as follows.[25] For the time being, let's take as given the values of the variance of the three patient level variables M ≡ ($\sigma_v^2, \sigma_w^2, \sigma_\varepsilon^2$). Given these, the average level of over- or under-optimism, $\bar{o}$, and ICDs value, $\bar{v}$, determine the average ICD use rate, while the variance of *perceived* skill ($\sigma_a^2 + \sigma_b^2$) determines the variance of ICD use across hospitals. Conditional mortality across hospitals is determined by M, $\bar{o}$, and the average value of the ICD, $\bar{v}$,, while the variance of one- and two-year conditional mortality helps us pin down the variance of true skill across hospitals/doctors and the relative variance of v and w.

The variance of μ, ($\sigma_v^2 + \sigma_w^2$), and the noise of the signals (σ_ε^2) is identified from the correlation between ICD use rates and conditional mortalities. Intuitively, α_i is identified by the

[23] The rate of ICD use among potentially appropriate patients, 18.5 percent, is derived from Al Khatab et al. (2012) based on their study of ICD use in a cohort of CHF patients; we assume that variation in this parameter is proportional to observed variation in population-based utilization, which is of course much lower. We know the mortality rate among those treated with an ICD, but we impute the mortality for those without an ICD (and the average treatment effect μ_j) used estimates from the largest randomized trial, which showed no impact after one year, and an approximately 2.5 percentage point reduction in mortality after 2 years (Bardy et al., 2005).

[24] All of these moments are computed over the period 2006-2013.

[25] In reality, some of the parameters affect more than one moment for example, the variances of all hospital level variables affect the variance of ICD use rates as well as the one- and two-year conditional mortalities.

correlation between ICD use and one-year mortality conditional on ICD implant; a high α reduces the sensitivity of the decision of implanting an ICD to the level of perceived skill, a_i^p. Therefore, those doctors with higher perceived skill (relative to actual skill) will not go as deep into the distribution of patients when α_i is low. For this reason, their marginal patient has a higher μ, leading to lower mortality rates. As a result, a higher α is associated with a lower correlation between ICD use and mortality conditional on having an ICD.

Table 3 reports the data and model-implied moments. The model does a good job of matching aggregate moments. The only target that the model misses is the average 2-year mortality. Table 4 reports the calibrated values for the aggregate parameters. Considering first the baseline estimates, on average for ICD use, hospitals are overly optimistic with regard to their skill, with the mean of b_i equal to 0.098. Variation in misperception is also substantial, with a standard deviation of 0.02, roughly 2.5 times the standard deviation of variation in skill (0.008). Pessimistic or under-confident hospitals are therefore rare in our sample. Perhaps not surprising, patient-level variation in individual outcomes are also estimated to be large, suggesting the presence of unmeasured health-related factors; this is necessary to match the relatively low cross-sectional dispersion of conditional mortality rates.

Hospital level parameters. Once we have calibrated the common parameters, we identify for each hospital and year the true skill and degree of misperception that produce the observed ICD use rate and conditional mortality. Given the aggregate parameters, equation (5) shows that the ICD use rate is fully determined by the perceived skill of the hospital. Therefore, we can identify perceived skill by inverting the expression for ICD use rate (5). Proposition 2 shows that, for a given perceived skill, the conditional mortality is decreasing in true skill. Therefore, we can invert equation (6) to identify the hospital/year true skill level.

We start by analyzing the identified levels of a_i and o_i; reassuringly, the distributions of hospital-level skill and overconfidence look approximately Normal, and the variance of a_i and b_i are close to the estimates of σ_a^2 and σ_b^2 identified in the calibration of the aggregate parameters.[26] We note in passing geographic differences in misperception; the regions with greater optimism or overconfidence are in the South (e.g. Texas), South-East, and the Great Lakes (Michigan, Indiana, Ohio).

[26] The variance of overconfidence is equal to 0.028 in the hospital-level identification vs. 0.025 in the aggregate calibration while the variance of skill is 0.0076 in the hospital-level identification vs. 0.0057 in the aggregate calibration.

To further understand the nature of variation in the hospital level estimates of skill and misperception we conduct a variance covariance decomposition. Specifically, let x_{it} be the estimate of *x* in hospital *i* and year *t*, for $x_{it} = \{a_{it}^p, a_{it}, b_{it}\}$. Then we can decompose the variance of x_{it} into the "within hospital" over time component, and the "between hospital" component (See Table 5).[27] For all three parameters, the variance of the within component is smaller than the variance of the between component. However, there is significant variation in the relative contribution of the within and between components across the three variables. For perceived skill, the variance of the within component is approximately half the variance of the between component. By contrast, for skill, the variance of the within component is less than one fourth the variance of the between component. That is, providers are more effective in learning about the bias in perceived skill, than they are with learning-by-doing, which would induce time-variation in true skill.

4.2 Analysis

Now that we have identified the key parameters of the model at both the aggregate and hospital level, we can return to our primary goal, which is to assess the relative contribution of skill and overconfidence for the evolution of ICD utilization and health outcomes.

4.2.1 Determinants of ICD use and conditional mortality. We start by exploring the empirical consequences of skill and misperception for ICD use and mortality. First, we decompose the fraction of observed variation attributed to differences across providers in skill, and differences in misperception. These simulations allow us to calculate the fraction of the variance in ICD use, or the variance of conditional mortality, attributable to variations in misperceptions across

[27] Specifically, let T_i denote the number of observations corresponding o hospital *i*, $\bar{x}_i$ the average of *x* in hospital *i*, and $\bar{\bar{x}}$ be the average of *x* across all hospitals. Then the within hospital *i* variance of *x* is

$$Var_i = \frac{\sum_t (x_{it} - \bar{x}_i)^2}{T_i}$$

The between hospital variance is defined as

$$Var_{be} = \frac{\sum_i (\bar{x}_i - \bar{\bar{x}})^2}{N_i}$$

Then variance of x_{it} can be expressed as:

$$Var(x_{it}) = \frac{\sum_i Var_i}{N_i} + Var_{be}$$

hospitals. Table 6 reports these calculations. The decomposition indicates that 72 percent of the hospital variation in ICD utilization is due to variation in misperception, with the remaining 28 percent because of variation in skill. Similarly, 55% of the hospital variation in (one-year) conditional mortality is due to variation in misperception, with the remainder owing to differences in skill. Therefore, we conclude that misperception is a major contributor to the observed variation in the adoption and diffusion of ICDs, and of equal importance in explaining variations in health outcomes.

Table 7 provides results from the structural model in which the mean level of misperception is set equal to zero (but with an unchanged variance), where the variance is set equal to zero, and in the final column, both the mean and variance are set equal to zero, essentially "turning off" misperception. Turning off misperception is predicted to reduce average ICD utilization rates from 17.5 percent to 13.9 percent, or a 21 percent reduction. Similarly, the extent of variation in ICD use is predicted to be cut to just 36 percent of the original standard deviation (from 4.46 percent to 1.64 percent), while mortality is predicted to decline by 7 percent (with a corresponding decline in the standard deviation of mortality of 22 percent). These counterfactual exercises provide more support for the importance of misperception in explaining variation in technology adoption and use.

4.2.2 Correlated skill and misperception

We generalize the baseline model by assuming that $b_i = \gamma * a_i + \varepsilon_i^o$. This assumption implies that the variance of overconfidence is $\sigma_b^2 = \gamma^2\sigma_a^2 + \sigma_{\varepsilon_i^b}^2$, where $\sigma_{\varepsilon_i^b}^2$ is the variance of ε_i^b, the variance of perceived skill is $\sigma_{a^p}^2 = (1+\gamma)^2\sigma_a^2 + \sigma_{\varepsilon_i^b}^2$, and the correlation between true skill and misperception is $corr(a_i, b_i) = \frac{\gamma}{\left[1+\frac{\sigma_{\varepsilon_i^b}^2}{\gamma^2\sigma_a^2}\right]^{1/2}}$.

The calibration of the aggregate parameters in this setting differs from the baseline in two respects. First, we need to calibrate γ. Second, instead of calibrating the variance of the extent of misperception, we need to calibrate the variance of ε_i^b. In terms of identification, the correlation between true skill and overconfidence affects the correlation between ICD use and conditional mortality across hospitals. For a given variance of perceived skill, a higher value of γ implies that those doctors that conduct more ICDs are also more skilled. Therefore, increasing γ while keeping

constant $\sigma_{a^p}^2$ should lead to a lower correlation between ICD use and conditional mortality across hospitals.

Table 4 provides evidence from a regression analysis that favors a positive correlation between skill (a_i) and misperception (b_i); the estimate of γ is 0.47. Given that risk-adjusted mortality is also higher in high-utilization regions, these estimates imply that in high-utilization hospitals, the adverse effects on clinical quality of overconfidence more than compensates for the somewhat higher skill levels. (In Column 3 of Table 4, we also find a positive covariance between skill (a_i) and the value to patients of being treated (v_i)). Note that these estimates imply that hospitals with greater skill are also overly optimistic, since the correlation between a_i and b is Still, our key structural estimates are largely unchanged regardless of the assumption about the covariance structure, so we continue to use our baseline estimates in the subsequent simulations.

4.2.3 Learning and the evolution of perceived skill. Next, we turn our attention to the time variation in overconfidence, and in perceived skill. In particular, we explore the ability of the learning model posed in section 2 to explain the evolution of perceived skill and conditional mortality. To this end, we first estimate the following econometric counterpart to the model dynamics characterized by equation (12):

$$a_{i,t+1}^p - a_{i,t}^p = \alpha_{0i} - \alpha_i^a * b_{i,t} + u_{i,t}. \qquad (14)$$

Comparing specification (14) with equation (12), we can interpret the coefficients α_{0i} and α_i^a. The intercept α_{0i} captures the average realization of $\alpha_{it}^a * \xi_{it}$ for hospital *i.* Unlike the learning coefficient in equation (12), we force the coefficient α_i^a to be constant over time though it can vary across hospitals. Therefore, α_i^a captures the average learning rate in hospital *i*.

Column I of Table 8 reports the median value of α_{0i} and α_i^a across hospitals. For comparison purposes, column II reports the estimates when the two parameters are restricted to be the same across hospitals.[28] The key finding is that, despite its simplicity, the learning model captures a significant portion of the annual variation in perceived skill. In the baseline specification (with hospital-specific coefficients), the learning model accounts for 50% of the variation in the

[28] The difference in the number of observations between both specifications is due to the fact that we require hospitals to have at least four observations to estimate the hospital-specific parameters.

change of perceived skill, while in the version where both the intercept and learning coefficients are restricted to be the same across hospitals, the R^2 still is 23%.

As predicted by our model, the median learning coefficient α_i^a is significantly positive and between zero and one. The median point estimate is 0.47 which implies that the variance of the noise in the signal about the doctor's skill is approximately the same as the variance of the prior of perceived skill. Interestingly, there is significant variation in the estimated learning coefficients. The standard deviation of α_i^a across hospitals is 0.6.[29] For comparison, the standard deviation of the intercept in (14) across hospitals is 0.13.

To gain further insights about the learning process, we estimate the cross-hospital association between the learning coefficient α_i^a and the overconfidence (b_{i06}) and skill (a_{i06}) in 2006, the initial year. That is, we run

$$\alpha_i^a = \gamma + \gamma_o b_{i06} + \gamma_a a_{i06} + \epsilon_i \qquad (15)$$

The fit of this specification is quite good with a R^2 of 0.57, which implies that initial overconfidence and skill account for much of the cross-hospital variation in how quickly they learn about their overconfidence. Both coefficients, γ_b and γ_a, are negative and significant.[30] The fact that hospitals with lower learning coefficient α_{ai} are both overly optimistic, and have better-than-average skill suggest that such hospitals have tighter priors about their skill. As a result, they perform more ICD procedures and, despite the greater number of signals, they learn more slowly about their true skill, cutting too gradually the rate of ICD use relative to other, more nimble hospitals.

We conclude our analysis of the drivers of the growth in perceived skill by studying the relationship between the hospital level intercept in (14), α_{0i}, and initial overconfidence and skill. To this end, we run a version of equation (15) but using α_{0i} as the dependent variable.[31] The coefficient on skill is negative (-0.1) and marginally significant, but the coefficient on overconfidence is positive. This suggests that the mean reversion in perceived skill is not a mechanical artifact that more overconfident hospitals have lower intercepts in (14). It really

[29] In 8% of the hospitals the estimate of the learning coefficient is negative, while in 21% it is greater than one.

[30] Their point estimates are -1.01, and -0.96 respectively with standard errors equal to 0.1 and 0.245.

[31] The R2 is lower (0.38).

follows from the fact that hospitals learn about their overconfidence and update their perceived skill.

The role of initial skill in the evolution of perceived skill is also of interest. Other things equal, hospitals with higher initial skill experience larger declines in perceived skill. This observation seems to downplay the role of learning by doing -- true skill improves as hospitals implant more ICDs -- in the evolution of perceived/true skill. To assess more directly the relevance of improvements in skill from implanting ICDs (learning-by-doing) is by estimating the following specification

$$a_{it+1} - a_{it} = \alpha + \delta_1 ICD_{it} + \delta_2 a_{it} + \epsilon_{it} \quad (16)$$

where ICD_{it} is the ICD use rate in hospital I in year t. The point estimates from this regression are $\hat{\delta}_1 = -0.127$, and $\hat{\delta}_2 = -0.04$, both statistically significant. Therefore, other things equal, we find no association between the secular trend in skill, and the lagged ICD use rate.

4.2.4 Out-of-sample implications of learning for ICD use and conditional mortality. Now that we have shown that our stylized learning model is able to capture quite accurately the evolution of perceived skill, we directly explore whether the decline in ICD use and conditional mortality between 2006 and 2013 can be a consequence of learning about misperception.

To investigate this hypothesis, we use the estimates of the learning model (column I of Table 8) to build a counterfactual measure of perceived skill due to learning. Then, we use our model to simulate the ICD use and conditional mortality levels in 2006 and 2013 for the counterfactual measures of perceived skill. Then we compare the evolution of the relevant moments under the counterfactual with those observed in the data. Table 9 presents the results from this exercise.

The first three columns of Table 9 report the moments for the ICD use rate, with the first two rows corresponding to 2006 and the second two rows corresponding to 2013. The similarity in the moments for 2006 between model and data is by construction since we take the initial estimated misperception (or overconfidence) and skill parameters for each hospital and those produce the same ICD use rates and conditional mortalities we observed in the data. However, the predictions of the model for 2013 are out of sample since the values of skill and misperception we use to generate them are produced by the learning model instead. Therefore, only if the learning dynamics capture the evolution of true and perceived skill that we see in the data, we can expect

the distributions of ICD use and conditional mortality across hospitals in 2013 to be similar to the data. The key finding from Table 9 is that the model's out-of-sample predictions for 2013 also provide a close match to the actual 2013 values. The learning model *fully* accounts for the observed 6.8 percentage point decline in the ICD use rate. The learning model also predicts the reduction in the dispersion in ICD usage rates across hospitals and the reduction in the correlation between skill and ICD use that we have observed in the data (decline from 0.94 to 0.895 in our out-of-sample prediction vs. from 0.94 to 0.905 in the data).

The second three columns present the moments for the conditional mortality rate. In the data, we observed a mild decline in the one-year conditional mortality rate from 12.8% to 12.0%. Our learning model fully accounts for this reduction in the conditional mortality. Furthermore, the cross-sectional dispersion of conditional mortality in 2013 and its correlation with true skill across hospitals is very similar in the model and in the counterfactual. Thus the evolution of physician beliefs about the efficacy of ICDs for this new population of CHF patients can explain both the sharp decline (or exnovation) in the use of ICDs during this period, as well as a more modest decline in conditional mortality rates.

5. Discussion

What drives the diffusion of new technologies? Research in economics has focused on factors primarily related to rates of return, whether because of input prices, differential factor productivity, or higher rates of return; the puzzle has often been why so many economic agents diffuse so slowly. In this paper, we allow for a different determinant of technological adoption and use -- misperception – where an individual's perception of their own skill and ability causes them to step up or scale back the use of a new technology, even when true skills do not correspond to their beliefs. In the case of a specific medical technology, implantable defibrillators (ICDs), these behavioral biases appear to be important quantitatively and explain otherwise puzzling empirical regularities.

How should the misperception parameter be interpreted? As noted above in Section 3, we cannot distinguish between a misperception about the population-based mean of the treatment from a misperception about the physician's own skills relative to others. Another explanation not previously considered is simply supplier-induced demand; the physician knows that the procedure is suboptimal for her patients, but she does the procedure anyway to make more money Chandra,

et al., 2011). However, the supplier-induced demand story is at odds with results from the calibrated learning model; why should the most "entrepreneurial" physicians who are cognizant of their bias towards net income over patient health be the ones most likely to scale back over time.[32] While we cannot rule out "outlier" physicians with a strong weight on revenue over patient welfare (as in the case of the Munster Indiana physician documented by Creswell, 2015), we view these as exceptions rather than the rule. More generally, even if we accept that "misperception" is a portmanteau that includes a number of alternative explanations, the major point of our estimation approach is that this is quantitatively important in explaining why there is so much variation in adoption and utilization of technological innovations.[33]

We acknowledge several limitations to the study, most of which are related to the specificity of the ICD technology. Many studies of learning-by-doing find improvements in mortality over time (Gong, 2017; Jovanovik and Nyarko, 1995; although see Huesch, 2009). The lack of strong progress in mortality that we observe may be explained by the long years of experience many physicians already have with implanting ICDs in other types of patients. The change in coverage for an entirely new population of patients (e.g., CHF patients) means it is less surprising that the learning that we observe in the data was with regard to appropriateness for patients, rather than technical skill *per se*.

Still, one might expect to see a sharp decline in ICD implantation rates for hospitals with the poorest mortality outcomes (as in Chandra et al., 2016). Yet there was little or no way for most physicians (or referral physicians) to observe their own skill, and to know whether their own risk-adjusted rates were above or below average. The SCD-HeFT trial could have provided a rough guideline for mortality (roughly 8 percent mortality in the first year), but the patient mix in the community was substantially older than in randomized trials, so community-level physicians had no benchmark against which they could compare outcomes in their patients compared to the trial physicians.

[32] One could argue that the entrepreneurial physicians were pressured to scale back because of pressure from CMS because CMS began to question out-of-guideline ICDs and threaten not to pay for them. However, we found (in unreported analysis) there was essentially no correlation between rates of ICD use and the fraction out-of-guideline, nor was there a strong correlation between conditional mortality and the fraction out-of-guideline.

[33] Another alternative explanation is potential demand-side factors – e.g., patient pressures to seek the procedure – that could also lead to systematically overusing or underusing the procedure. While patient preferences are likely to be important at the individual level, there is little evidence that such preferences can explain regional differences (Cutler et al., 2019).

More problematic is that ICDs appear to have been less successful than expected (e.g., MacMurray, 2016) in part because of their side-effects; thus physicians may have viewed the initial trials as a particularly favorable signal, but they updated views over time. This pattern of enthusiasm followed by disappointment is not uncommon for medical procedures; as Jupiter and Burke (2013) have written:

> Artelon® arthroplasty, thermal shrinkage, Vioxx®, metal-on-metal hip arthroplasty, and Infuse® bone grafting in the spine—all had come onto the "market" with enthusiastic reports only to fall from grace to unhappy outcomes, permanent disabilities, and malpractice litigation. (p. 249).

By contrast, other innovations have begun with much lower expectations, but ended up delivering large patient benefits. For example, Currie, MacLeod, and Van Parys (2015) found that cardiologists who were more aggressive than then-current guidelines for percutaneous coronary interventions (PCI, or stenting) gained better results; in this case physicians gradually updated their priors that stents were more productive than expected. We need not take a stand on the average level of misperception, whether positive or negative (or zero); the key finding is that at least for ICDs, the variance of misperception is large and persistent, thus explaining why we observe such a large variance in both adoption and the use of this technology.

Still, how generalizable is the case of ICDs to technology outside of health care? The result that physicians overestimate their own skill level is certainly consistent with other data from laboratory experiments in which hypothetical entrepreneurs are overconfident about their own ability and enter into markets or games where failure is likely (Camerer and Lovallo, 1999). And a pattern of overconfidence is common across non-physicians, as for example with regard to individual assessment of one's own driving skills (Svenson, 1981). Further data and case studies are clearly required, but despite these caveats, it appears there is a first-order role for misperception in both explaining the wide variations in the adoption and diffusion of new technologies, and in attenuating the aggregate productivity of new technologies.

References

Ahmed, A., W.S. Aronow, and J.L. Fleg.. "Higher New York Heart Association Classes and increased mortality and hospitalization in heart failure patients with preserved left ventricular function." *American Heart Journal* 151(2): 444-450, 2006.

Al-Khatib, S. M., A. Hellkamp, J. Curtis, D. Mark, E. Peterson, G. D. Sanders, P. A. Heidenreich, A. F. Hernandez, L. H. Curtis, and S. Hammill. Non–evidence-based ICD implantations in the united states. *JAMA*, 305(1):43–49, 2011.

Al-Khatib, S. M., Hellkamp, A. S., Hernandez, A. F., Fonarow, G. C., Thomas, K. L., Al-Khalidi, H. R., & Peterson, E. D. Trends in use of implantable cardioverter-defibrillator therapy among patients hospitalized for heart failure. *Circulation*, 125(9), 1094-1101, 2012.

Arni, P., D. Dragone, L. Goette, and N.R. Ziebarth, "Biased health perceptions and risky health behaviors – Theory and evidence," Working paper, July 2018.

Barber, B.M. and T. Odean. Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics*, pages 261–292, 2001.

Bardy, G. H., Lee, K. L., Mark, D. B., Poole, J. E., Packer, D. L., Boineau, R., ... & McNulty, S. E. Amiodarone or an implantable cardioverter–defibrillator for congestive heart failure. *New England J Medicine*, 352(3), 225-237, 2005.

Basu S. and D. Weil. Appropriate technology and growth. *The Quarterly Journal of Economics*, 113(4):1025–1054, 1998.

Bekelis, K., Skinner, J., Gottlieb, D., and Goodney, P., De-adoption and exnovation in the use of carotid revascularization: retrospective cohort study. *BMJ*, 359, j4695, 2017.

Berner, E. S., and Graber, M. L. Overconfidence as a cause of diagnostic error in medicine. The American journal of medicine, 121(5), S2-S23, 2008.

Camerer, C. and Lovallo, D. Overconfidence and excess entry: An experimental approach. *American Economic Review*, 89(1), 306-318, 1999.

Caselli, F., I. Coleman, and W. John. The world technology frontier. *American Economic Review*, 96(3):499–522, 2006.

Chandra, A., A. Finkelstein, A. Sacarny, and C. Syverson. Healthcare exceptionalism? productivity and allocation in the U.S. healthcare sector. *American Economic Review*, 2016.

Chandra, A. and D. O. Staiger. Productivity spillovers in healthcare: evidence from the treatment of heart attacks. *The Journal of Political Economy*, 115:103, 2007.

Chandra, A., D. Cutler, and Z. Song. Who ordered that? The economics of treatment choices in medical care. In M.V. Pauly, T.G. McGuire, and P. Barros (eds.) Handbook of Health Economics Volume 2, Elsevier, 2011.

Clyde, A.T., L. Bockstedt, J. A. Farkas, and C. Jackson. Experience with Medicare's new technology add-on payment program. *Health Affairs*, 27(6):1632–1641, 2008.

Comin, D., and B. Hobijn. "Cross-country Technological Adoption: Making the Theories Face the Facts." *Journal of Monetary Economics* (January 2004).

Comin D., and B. Hobijn. "Implementing Technology." November 2007

Comin, D., and B. Hobijn. "An Exploration of Technology Diffusion." *American Economic*

Review 100, no. 5 (December 2010): 2031-59

Comin, D., and M. Mestieri. "If Technology Has Arrived Everywhere, Why Has Income Diverged?" *American Economic Journal: Macroeconomics*, 10 (3): 137-78, July 2018

Conley, T.G. and C. R. Udry. Learning about a new technology: Pineapple in Ghana. *American Economic Review*, pages 35–69, 2010.

Creswell, J. A small Indiana town scarred by a trusted doctor. *The New York Times*, October 17 2015.

Currie, J., and W. B. MacLeod. Diagnosis and unnecessary procedure use: Evidence from C-section. National Bureau of Economic Research, 2013.

Currie, J., W. B. MacLeod, and J. Van Parys. Physician practice style and patient health outcomes: The case of heart attacks. *Journal of Health Economics*, 2016.

Cutler, D., J. Skinner, A.D. Stern, and D. Wennberg, "Physician beliefs and patient preferences: A new look at regional variation in health care spending," *AEJ: Economic Policy*, 2019; NBER Working Paper No. 19320.

Duffy, S. Q., and Farley, D. E. The protracted demise of medical technology: The case of intermittent positive pressure breathing. *Medical Care*, 718-736, 1992.

Duflo, E., M. Kremer, and J. Robinson, Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya. *The American Economic Review*, 101(6), 2350-2390, 2011.

Finkelstein, A., Gentzkow, M., Hull, P., and H.Williams, Adjusting risk adjustment–Accounting for variation in diagnostic intensity. *New England J Medicine*, 376(7), 608, 2017.

Foster, A.D., and M. R. Rosenzweig. Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy*, 1176–1209, 1995.

Friedrich, E.B., and M. Böhm. Management of end stage heart failure. *Heart*, 93(5):626–631, 2007.

Gong, Q. Physician learning and treatment choices: Evidence from brain aneurysms, working paper, University of Pennsylvania, 2017.

Griliches, Z. Hybrid corn: An exploration in the economics of technological change. *Econometrica*, pages 501–522, 1957.

Heusch, M.D. Learning by doing, scale effects or neither? Cardiac surgeons after residency, Health Services Research, 44(6): 1960-82, 2009.

Jovanovic, B., and Y. Nyarko. A Bayesian learning model fitted to a variety of empirical learning curves. *Brookings Papers on Economic Activity. Microeconomics*, 1995:247–305, 1995.

Jovanovic, B., and Y. Nyarko. Learning by doing and the choice of technology. *Econometrica*, pages 1299–1310, 1996.

Jupiter, J., and D. Burke. "Scott's parabola and the rise of the medical-industrial complex." *Hand*: 249-252, 2013.

Kadish, A., A. Dyer, J. P. Daubert, R. Quigg, N. M. Estes, K. P. Anderson, H. Calkins, D. Hoch, J. Goldberger, A. Shalaby, et al. Prophylactic defibrillator implantation in patients with nonischemic dilated cardiomyopathy. *New England J of Medicine*, 350(21):2151–2158, 2004.

Kolata, Gina, "For patients with heart failure, little guidance as death nears," *New York Times*, November 7, 2017.

Kolata, Gina, "Tiny device Is a 'huge advance' for treatment of severe heart failure," *New York Times*, September 23, 2018.

Levy, W.C., I. Goldenberg, and A. Moss. Benefits of ICD therapy: Connecting treatment decisions to individualized ICD risk estimates within MADIT-II using the Seattle proportional risk model. *Journal of the American College of Cardiology*, 65(10 S), 2015.

Lindvall, C., N. A. Chatterjee, Y. Chang, B. Chernack, V. A. Jackson, J. P. Singh, and J. P. Metlay. National trends in the utilization of cardiac resynchronization therapy with or without implantable-cardioverter defibrillator. *Circulation*, pages 115, 2015.

Lucas, R.E., Jr. On the size distribution of business firms. *The Bell Journal of Economics*, pages 508–523, 1978.

Malmendier, U., and G. Tate. Does overconfidence affect corporate investment? CEO overconfidence measures revisited. *European Financial Management*, 11(5):649–659, 2005.

Manski, C.F. Diagnostic testing and treatment under ambiguity: Using decision analysis to inform clinical practice. *Proceedings of the National Academy of Sciences*, 110(6):2064–2069, 2013.

Matchett, M., S. F. Sears, G. Hazelton, K. Kirian, E. Wilson, and R. Nekkanti. The implantable cardioverter defibrillator: its history, current psychological impact and future. *Expert review of medical devices*, 6(1):43–50, 2009.

McMurray, J. J. The ICD in heart failure-time for a rethink? *New England J of Medicine*, 375(13): 1283-1284, 2016.

Moen, E. L., Bynum, J. P., Austin, A. M., Skinner, J., Chakraborti, G., and O'Malley, A. J. Assessing variation in implantable cardioverter defibrillator therapy guideline adherence with physician and hospital patient-sharing networks. *Medical Care* 56(4): 350-357, 2018.

Moore, D. A., and Healy, P. J. The trouble with overconfidence. *Psychological Review*, 115(2), 502-517, 2008.

Moss, A.J., W. Zareba, W. J. Hall, H. Klein, D. J. Wilber, D. S. Cannom, J. P. Daubert, S. L. Higgins, M. W. Brown, and M. L. Andrews. Prophylactic implantation of a defibrillator in patients with myocardial infarction and reduced ejection fraction. *New England Journal of Medicine*, 346(12):877–883, 2002.

O'Mahony, C., P. D. Lambiase, G. Quarta, M. Cardona, M. Calcagnino, K. Tsovolas, S. Al-Shaikh, S. M. Rahman, S. Arnous, S. Jones, et al. The long-term survival and the risks and benefits of implantable car- dioverter defibrillators in patients with hypertrophic cardiomyopathy. *Heart*, 98(2):116–125, 2012.

O'Neill, H.M., R. W. Pouder, and A. K. Buchholtz. Patterns in the diffusion of strategies across organizations: Insights from the innovation diffusion literature. *Academy of Management Review*, 23(1):98–114, 1998.

Pavcnik, N., Trade liberalization, exit, and productivity improvements: Evidence from Chilean plants. *Review of Economic Studies*, 69(1), 245-276, 2002.

Poole, J.E.. Present guidelines for device implantation clinical considerations and clinical challenges from pacing, implantable cardiac defibrillator, and cardiac resynchronization therapy.

Circulation, 129(3):383–394, 2014.

Roger V.L. Epidemiology of heart failure. *Circulation Research*. 113(6):646-659, 2013.

Rogers E.M. Diffusion of innovations. Simon and Schuster; 2004.

Roth, G.A., J. E. Poole, R. Zaha, W. Zhou, J. Skinner, and N. E. Morden. Use of guideline-directed medications for heart failure before cardioverter- defibrillator implantation. *Journal of the American College of Cardiology*, 67(9):1062–1069, 2016.

Sanders, G.D., M. A. Hlatky, and D. K. Owens. Cost-effectiveness of implantable cardioverter–defibrillators. *New England J Medicine,* 353(14):1471–1480, 2005.

Schröder, J.H., J. Hugosson, M. J. Roobol, T. L. Tammela, S. Ciatto, V. Nelen, M. Kwiatkowski, M. Lujan, H. Lilja, M. Zappa, et al. Prostate-cancer mortality at 11 years of follow-up. *New England J Medicine*, 366(11):981–990, 2012.

Skinner J., and D. Staiger. Technological diffusion from hybrid corn to beta blockers. *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches.* University of Chicago Press and NBER, 2007.

Skinner, J., and D. Staiger. Skinner, J., & Staiger, D. Technology diffusion and productivity growth in health care. Review of Economics and Statistics, 97(5), 951-964, 2015.

Song, Y., Skinner, J., Bynum, J., Sutherland, J., Wennberg, J. E., & Fisher, E. S. Regional variations in diagnostic practices. *New England J Medicine*, 363(1), 45-53, 2010.

Suri, T. Selection and comparative advantage in technology adoption. *Econometrica*, 79(1):159–209, 2011.

Svenson, Ola. "Are we all less risky and more skillful than our fellow drivers?" *Acta Psychologica*, 47(2), pp. 143-48, 1981.

van Rees, J.B. *Implantable cardioverter defibrillators: translating evidence from randomized clinical trials to routine clinical practice.* PhD thesis, Department Cardiology, Faculty of Medicine/Leiden University Medical Center (LUMC), Leiden University, 2014.

Wennberg, D.E., F. Lucas, J. D. Birkmeyer, C. E. Bredenberg, and E. S. Fisher. Variation in carotid endarterectomy mortality in the medicare population: trial hospitals, volume, and patient characteristics. *JAMA*, 279(16):1278–1281, 1998.

Wennberg, D.E., Sharp, S.M., Bevan, et al., ”A population health approach to reducing observational intensity bias in health risk adjustment: cross sectional analysis of insurance claims, *BMJ* 348, p. 2392, 2014.

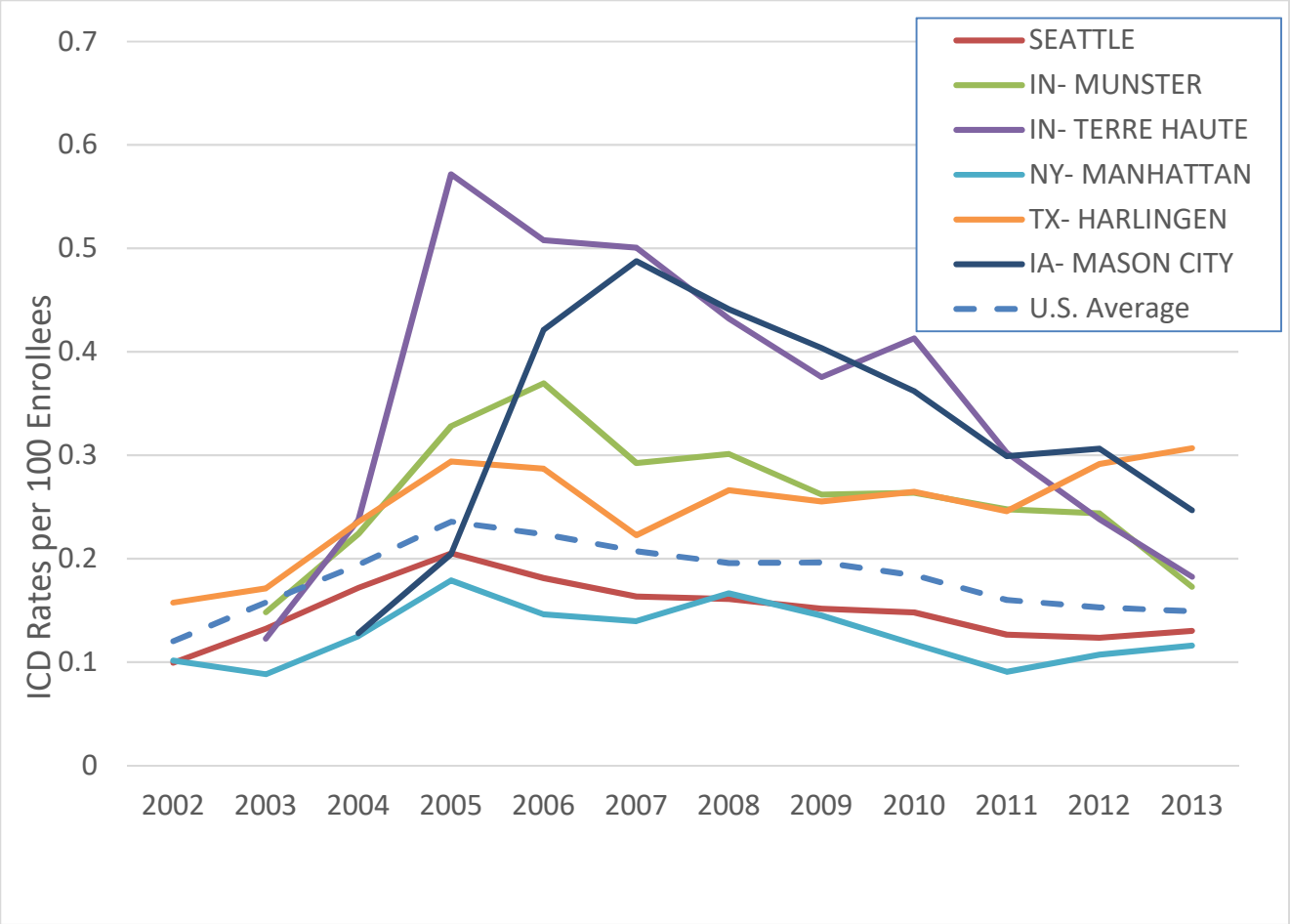

Figure 1: Risk-Adjusted Rates of ICD use per 100 Medicare Enrollees for Selected Hospital Referral Regions, 2002-13.

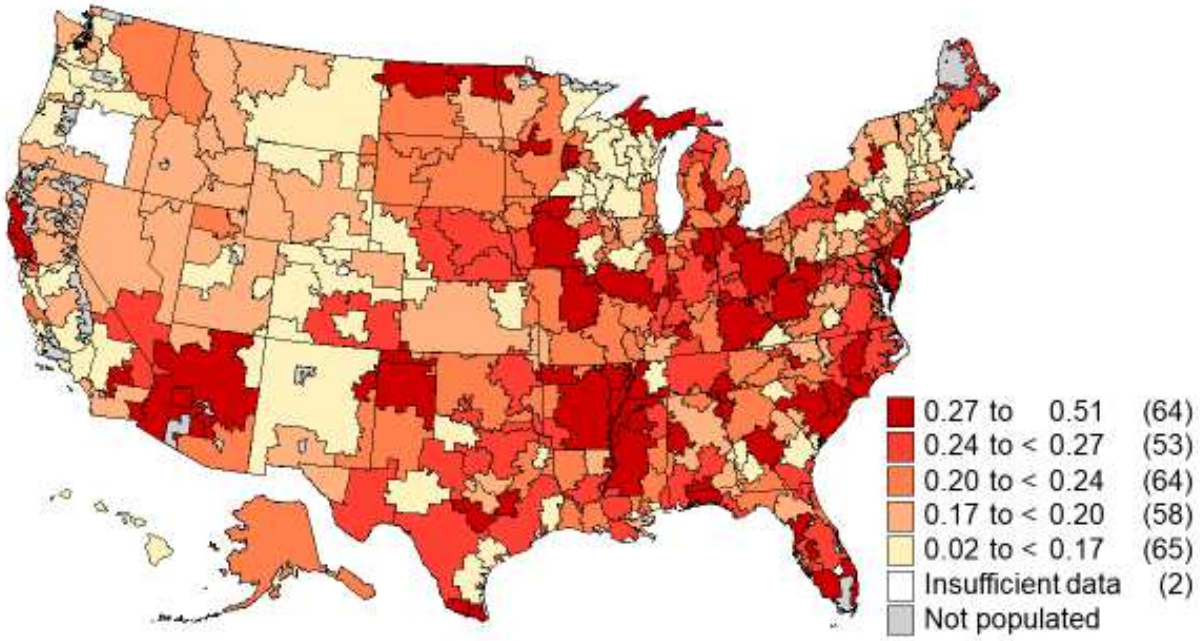

Figure 2. Implantable Cardioverter Defibrillator (ICD) rates per 100 Medicare enrollees, 2006. Risk adjusted for poverty, income, Medicaid dual-eligible, age, sex, race, county-level smoking, diabetes, obesity, and individual CHF diagnosis in the over-65 Medicare Fee-for-service population.

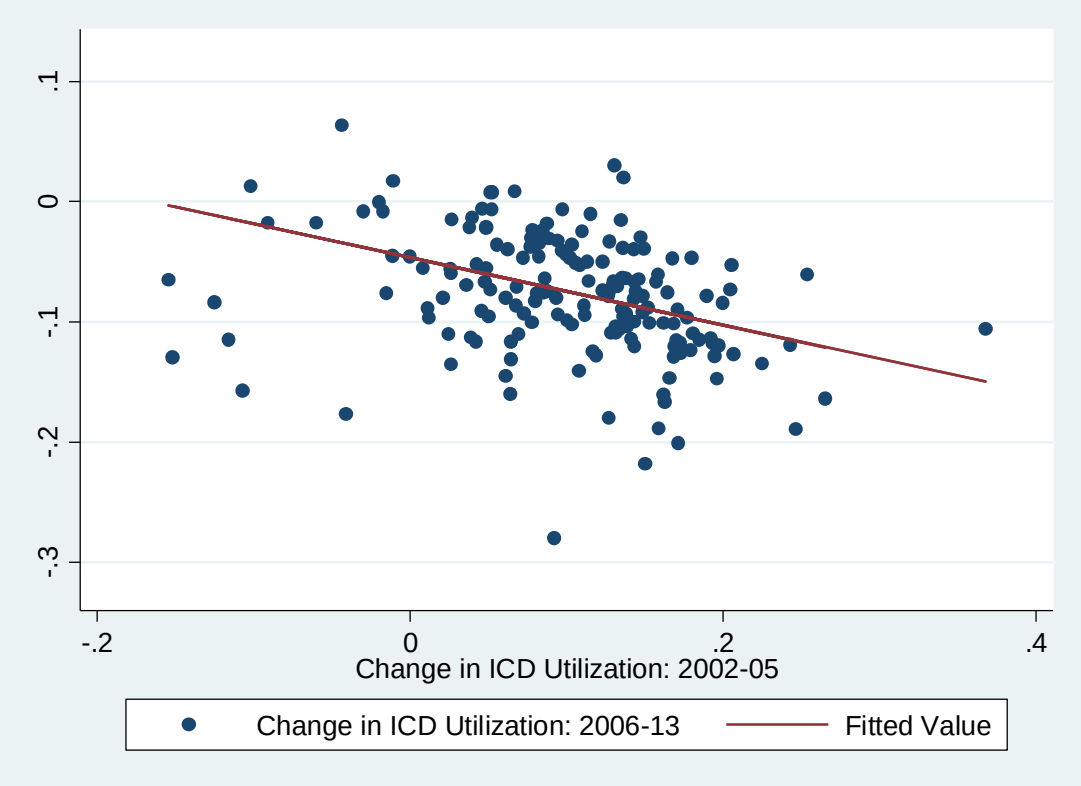

Figure 3: Correlation between 2002-05 and 2006-13 increases in Risk-Adjusted ICD Utilization Rates, at the HRR level (Rates per 100 Medicare enrollees for all types of ICDs).

Figure 4a: Assigning regional ICD use rates to hospitals

Figure 4b: Assigning hospital-level mortality to HRRs

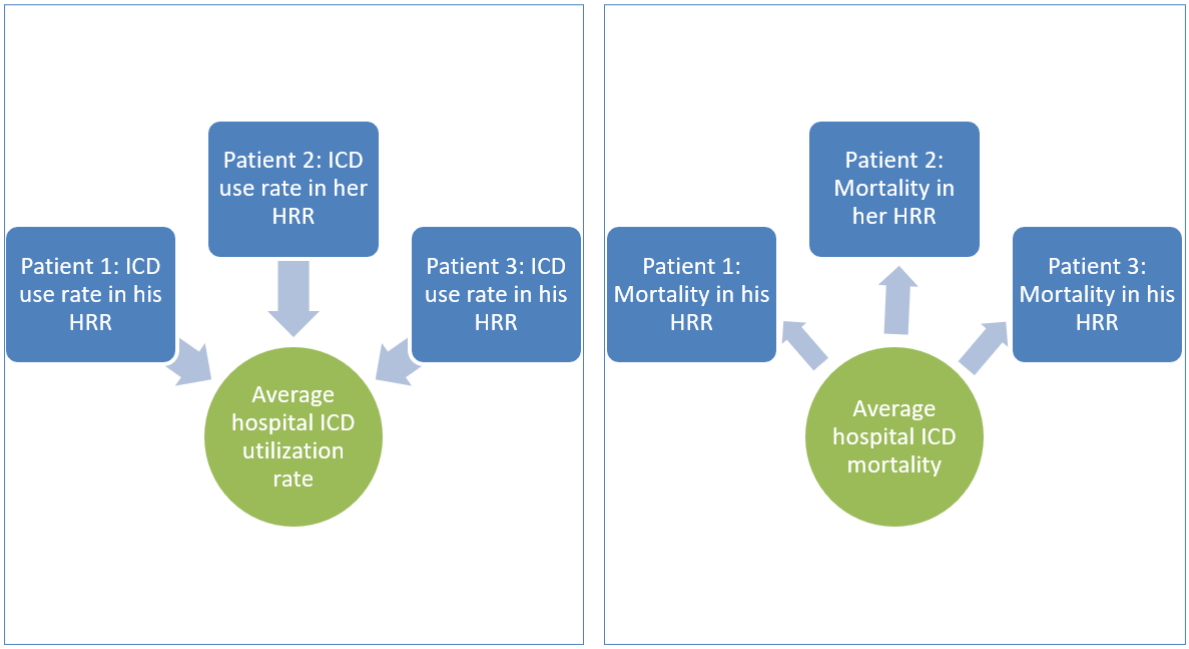

Figure 4: Schematic to show how ICD utilization rates are assigned to hospitals, and how hospital mortality rates are assigned to hospital referral regions (HRRs).

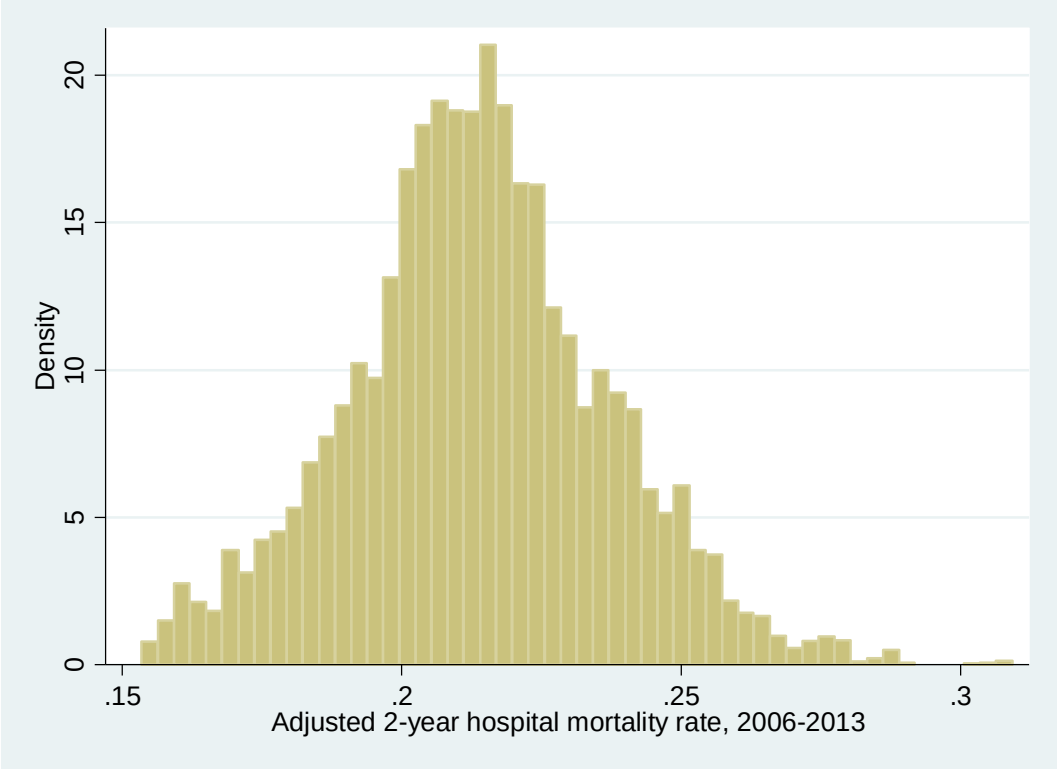

Figure 5: Distribution of Risk-adjusted Random-Effects 2-Year Mortality by Hospital: 2006-13

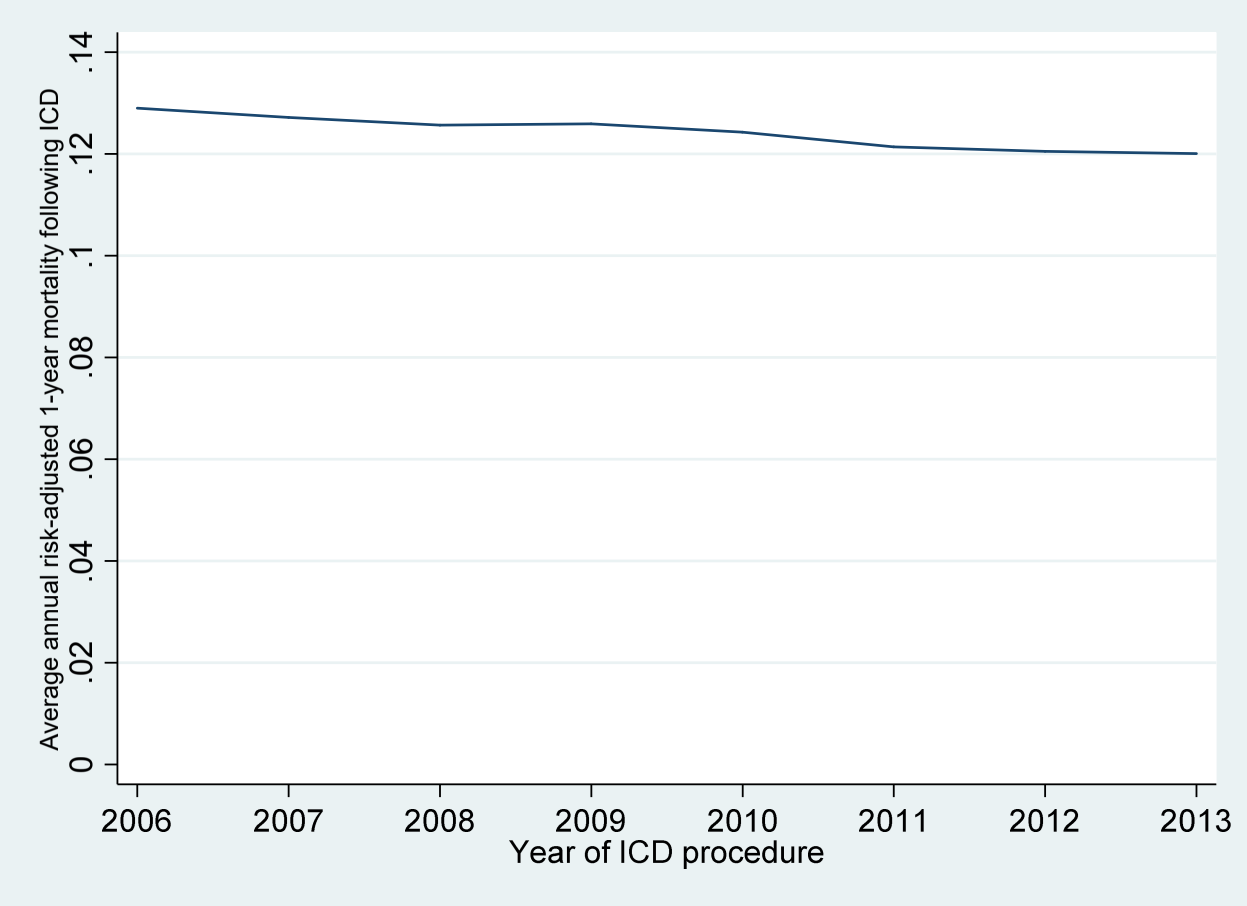

Figure 6: Evolution of 1-Year Mortality, 2006-13

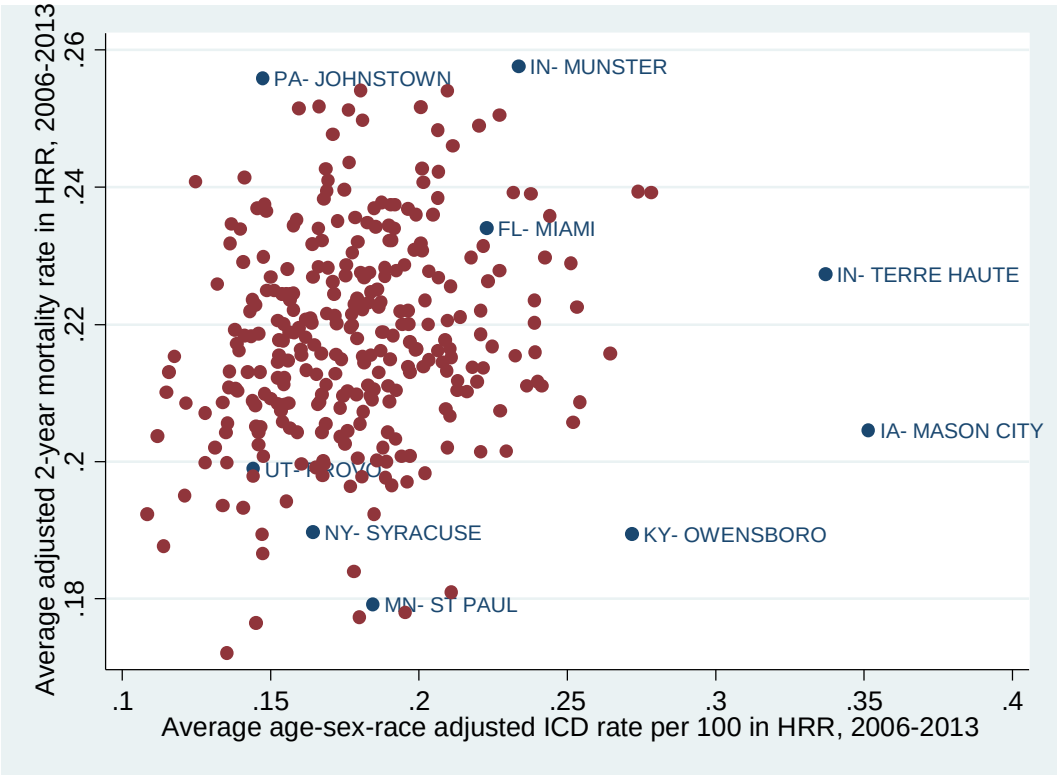

Figure 7: Correlation Between Average ICD Utilization (2006-13) and 2-Year Risk-adjusted Mortality

Table 1: Summary Statistics for ICD Registry Data

Variable	Mean	Standard Deviation
2-Year Mortality: 2006-13	0.218	0.413
2-Year Mortality: 2006	0.219	0.414
2-Year Mortality: 2013	0.216	0.411
1-Year Mortality: 2006-13	0.123	0.328
1-Year Mortality: 2006	0.122	0.328
1-Year Mortality: 2012	0.118	0.323
Fraction Inappropriate	0.098	0.297
Ejection Fraction (Percentage)	25.766	7.319
Fraction with EF > 35%	0.034	0.182
Fraction Class I	0.029	0.169
Fraction Class IV	0.043	0.202
Age	74.897	6.248
Previous cardiac arrest	0.020	0.142
Family history: Sudden death	0.030	0.171
Ventricular tacchycardia	0.225	0.418
Non-ischemic dilated cardiomyopathy	0.320	0.467
Ischemic heart disease	0.696	0.460
Previous myocardial infarction	0.548	0.498
Previous CABG	0.395	0.489
Previous PCI	0.345	0.475
Electrophysiology study	0.083	0.276
VT indication (ES study)	0.021	0.143
Female	0.282	0.450
Black	0.101	0.301
Hispanic (Medicare)	0.052	0.222
Other race	0.025	0.157
Hispanic ethnicity (Registry)	0.051	0.219

Table 2: Regression Coefficients for OLS, Random, and Fixed Effects Models: Two-Year Mortality

VARIABLES	(1) OLS	(2) OLS	(3) Random Effect	(4) Random Effect	(5) Fixed Effect	(6) Fixed Effect
HRR-level ICD Rate	0.0765***	0.103***	0.128***	0.0952***	0.0445	0.0531
	(0.0291)	(0.0290)	(0.0263)	(0.0276)	(0.0470)	(0.0445)
Ln(volume)		-0.0133***		-0.0130***		-0.00925***
		(0.00139)		(0.00130)		(0.00277)
HRR-level Rx Rate		-0.122***		-0.126***		-0.0986
		(0.0186)		(0.0183)		(0.0987)
Observations	254,237	253,247	254,237	254,237	254,237	253,596
R-squared	0.051	0.051			0.057	0.057
Number of Groups			1,548	1,542		

Note: Covariates included in all regressions – see Appendix Tables A.1 (OLS), A.2 (Random Effects), and A.3 (Fixed Effects) for full sets of estimates. Robust standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 3: Estimated Moments and Model Fit

	Data	Baseline
Average use of ICDs among candidates for an ICD	0.185	0.186
Standard deviation (risk-adjusted) of the use of ICDs across hospitals	0.047	0.042
One-year mortality rate conditional on an ICD	0.122	0.112
Two-year mortality rate conditional on an ICD	0.218	0.307
Standard deviation of one-year mortality rate conditional on ICD across hospitals	0.023	0.020
Standard deviation of two-year mortality rate conditional on ICD across hospitals	0.031	0.036
Correlation between ICD use and one-year mortality rate conditional on ICD use, across hospitals	0.148	0.144
Correlation between ICD use and two-year mortality rate conditional on ICD use, across hospitals	0.111	0.132

Table 4: Parameter Estimates with Sensitivity Tests when a, o, and v are correlated

	Baseline	a_i and o_i Correlated	a_i and v_i Correlated
$\sigma^2_{w_j}$	1.090	1.075	1.283
$\sigma^2_{o_i}$	0.020	0.022	0.022
$\sigma^2_{a_i}$	0.008	0.008	0.009
$\sigma^2_{\varepsilon_j}$	1.163	1.062	1.126
$\sigma^2_{v_j}$	0.764	0.879	1.071
$\bar{o}$	0.098	0.105	0.116
$\bar{v} - \bar{\omega} - \bar{a}$	-1.066	-1.126	-1.043
γ		0.473	
λ			0.251
Gap in Moments	0.072	0.051	0.055

Table 5: Variance Covariance Decomposition of Hospital-Level Parameters

	a_p	a	o
Within Component	0.0168	0.0015	0.0086
Between Component	0.0215	0.0049	0.0139
Total Variance	0.0383	0.0064	0.0225

Table 6: Contribution of Skill and Overconfidence to ICD use and Conditional Mortality

	ICD Use	Conditional Mortality
Skill	27.6%	44.5%
Overconfidence	72.4%	55.5%

Table 7: Counterfactual Policy Experiments

	Data	Mean o_i =0	STD(o_i)=0	Mean o_i =0 & STD(o_i)=0
Average ICD rate	0.1753	0.1422	0.1562	0.01387
Std(ICD rate)	0.0446	0.0417	0.0178	0.0164
Average Conditional Mortality	0.1244	0.116	0.1242	0.116
Std(Conditional Mortality)	0.0149	0.0144	0.0124	0.0116
Correlation (ICD rate, Conditional Mortality)	0.219	0.324	-0.9943	-0.9893

Table 8: Regression Results from the Learning Model

Learning Model		
	I	II
α_0	-0.023*	0.0281
	(0.0013)	(0.0113)
α_a	0.471*	0.1953
	(0.007)	(0.0067)
N	7776	8359
R^2	0.49	0.23

Table 9: Time-Variation Induced by Learning

		ICD use			Conditional Mortality		
		Mean	Std	Corr with a_i	Mean	Std	Corr with a_i
2006	Data	0.21	0.053	0.94	0.128	0.016	0.65
	Model	0.21	0.053	0.94	0.129	0.018	0.695
2013	Data	0.142	0.033	0.905	0.1205	0.0152	0.644
	Model	0.142	0.032	0.895	0.1198	0.0174	0.708

Appendix

Expression (4) defines the diffusion of ICD for a doctor/hospital with a given perceived skill. To compute the aggregate diffusion of ICDs we just need to compute the expectation of (4) over the initial distribution of perceived skills across hospitals. Formally, the diffusion of ICD in population is given by

$$\Pr(ICD) = \int_{-\infty}^{\infty} f_{a^p}(q) \Pr(\ ICD = 1|q) dq = \int_{-\infty}^{\infty} f_{a^p}(q) \left(\int_{\bar{s}(q)}^{\infty} f(s)\, ds \right) dq \qquad \text{(A.1)}$$

where $f_{a^p}(.)$ is the distribution of perceived skill in population.

Similarly, we can compute the standard deviation of the use of ICD's across hospitals as

$$Std(ICD_i) = SQRT\left[\int_{-\infty}^{\infty} f_{a^p}(q) (\Pr(\ ICD = 1|q) - \Pr(ICD))^2 dq \right] \qquad \text{(A.2)}$$

The mortality rate conditional on ICD implant is

$$Pr\left(v_j + a_i \leq \underline{\kappa} \middle| ICD = 1\right) = \frac{\Pr\left(v_j \leq \underline{\kappa} - a_i \cap\ ICD = 1\right)}{\Pr(\ ICD)}$$

$$= \frac{\int_{-\infty}^{\infty} f_{a,a^p}(q,q^p) \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\varepsilon}(\varepsilon') f_w(w') \left(\int_{\bar{s}(q^p)-\varepsilon'-w'}^{\underline{\kappa}-q} f_{\mu}(\mu') d\mu' \right) d\varepsilon' dw' \right) dq dq^p}{\Pr(\ ICD)} \qquad \text{(A.3)}$$

where $f_{a,a^p}(.,.)$ is the joint distribution of the duple skill, and perceived skill in population.

The standard deviation of mortality rates across hospitals is computed as

$$Std(Mort_i) = SQRT\left[\int_{-\infty}^{\infty} f_{a,a^p}(q,q^p) \left(Pr\left(v_j + a_i \leq \underline{\kappa} \middle| ICD = 1, q^p, q\right) - Pr\left(v_j + a_i \leq \underline{\kappa} \middle| ICD = 1\right) \right)^2 dq\, dq^p \right] \qquad \text{(A.4)}$$

Finally, the correlation between mortality and ICD use across hospitals is computed as

$$\rho_{Mort_i, ICD_i} = \frac{Covar(ICD_i, Mort_i)}{Std(ICD_i) Std(Mort_i)} \qquad \text{(A.5)}$$

Table A.1: Mortality (One & Two Years) OLS Regression

VARIABLES	(1) death1yr	(2) death1yr	(3) death1yr	(4) death2yr	(5) death2yr	(6) death2yr
HRR-level ICD Rate	0.0711***	0.105***	0.0849***	0.0765***	0.130***	0.103***
	(0.0225)	(0.0229)	(0.0226)	(0.0291)	(0.0297)	(0.0290)
Ln(volume)		-0.00829***	-0.00826***		-0.0133***	-0.0133***
		(0.00108)	(0.00108)		(0.00140)	(0.00139)
HRR-level Rx Rate			-0.0921***			-0.122***
			(0.0140)			(0.0186)
Ejection Fraction (EF) <20%	-0.00345***	-0.00346***	-0.00348***	-0.00442***	-0.00444***	-0.00448***
	(0.000382)	(0.000381)	(0.000381)	(0.000473)	(0.000472)	(0.000471)
EF 20-25%	-0.00459***	-0.00459***	-0.00457***	-0.00548***	-0.00548***	-0.00546***
	(0.000407)	(0.000407)	(0.000407)	(0.000502)	(0.000501)	(0.000502)
EF 25-30%	-0.00224***	-0.00227***	-0.00227***	-0.00388***	-0.00393***	-0.00393***
	(0.000355)	(0.000355)	(0.000356)	(0.000449)	(0.000449)	(0.000451)
EF 30-35%	-0.000740*	-0.000734*	-0.000732*	-0.00109**	-0.00108**	-0.00106*
	(0.000415)	(0.000414)	(0.000414)	(0.000555)	(0.000552)	(0.000552)
EF > 35%	0.00151***	0.00150***	0.00149***	0.00197***	0.00195***	0.00191***
	(0.000324)	(0.000323)	(0.000323)	(0.000398)	(0.000398)	(0.000399)
EF Missing	0.0226***	0.0213***	0.0211***	0.0319***	0.0297***	0.0278***
	(0.00719)	(0.00720)	(0.00725)	(0.00967)	(0.00966)	(0.00959)
NY Heart Assoc. Class II	0.00293	0.00245	0.00160	0.00583	0.00506	0.00399
	(0.00342)	(0.00339)	(0.00338)	(0.00489)	(0.00482)	(0.00481)
NY Heart Assoc. Class III	0.0480***	0.0476***	0.0465***	0.0707***	0.0701***	0.0687***
	(0.00346)	(0.00341)	(0.00340)	(0.00490)	(0.00480)	(0.00480)
NY Heart Assoc. Class IV	0.154***	0.153***	0.151***	0.191***	0.188***	0.187***
	(0.00583)	(0.00583)	(0.00584)	(0.00728)	(0.00722)	(0.00724)
NY Heart Assoc. Class missing	0.0549***	0.0520***	0.0516***	0.0928***	0.0883***	0.0880***
	(0.0106)	(0.0105)	(0.0106)	(0.0131)	(0.0129)	(0.0131)
Age 70-74	0.0150***	0.0151***	0.0153***	0.0275***	0.0277***	0.0276***
	(0.00165)	(0.00165)	(0.00165)	(0.00206)	(0.00205)	(0.00205)
Age 75-79	0.0367***	0.0368***	0.0368***	0.0648***	0.0651***	0.0651***
	(0.00181)	(0.00181)	(0.00182)	(0.00228)	(0.00227)	(0.00227)
Age 80-84	0.0633***	0.0635***	0.0636***	0.110***	0.110***	0.110***
	(0.00213)	(0.00213)	(0.00213)	(0.00258)	(0.00258)	(0.00258)
Age 85-89	0.103***	0.103***	0.102***	0.176***	0.176***	0.175***
	(0.00333)	(0.00334)	(0.00334)	(0.00410)	(0.00413)	(0.00410)
Age 90+	0.184***	0.184***	0.183***	0.276***	0.276***	0.276***
	(0.0107)	(0.0107)	(0.0107)	(0.0124)	(0.0124)	(0.0123)
Previous cardiac arrest	0.0578***	0.0569***	0.0569***	0.0578***	0.0564***	0.0567***
	(0.00549)	(0.00549)	(0.00550)	(0.00631)	(0.00632)	(0.00633)
Family history sudden arrest	-0.0116***	-0.0119***	-0.0116***	-0.0190***	-0.0195***	-0.0192***
	(0.00401)	(0.00394)	(0.00392)	(0.00492)	(0.00483)	(0.00481)
Ventricular tacchycardia	0.0447***	0.0449***	0.0447***	0.0570***	0.0573***	0.0570***
	(0.00193)	(0.00193)	(0.00193)	(0.00230)	(0.00230)	(0.00230)

Non-ischemic dilated cardiomyopathy	-0.0218***	-0.0211***	-0.0209***	-0.0320***	-0.0310***	-0.0306***
	(0.00247)	(0.00248)	(0.00247)	(0.00316)	(0.00317)	(0.00316)
Ischemic heart disease	0.0167***	0.0172***	0.0171***	0.0262***	0.0269***	0.0269***
	(0.00271)	(0.00271)	(0.00271)	(0.00332)	(0.00333)	(0.00333)
Previous myocardial infarction	0.00764***	0.00796***	0.00805***	0.0122***	0.0127***	0.0129***
	(0.00171)	(0.00171)	(0.00171)	(0.00223)	(0.00222)	(0.00222)
Previous CABG	0.00882***	0.00891***	0.00867***	0.0204***	0.0206***	0.0203***
	(0.00165)	(0.00165)	(0.00165)	(0.00201)	(0.00200)	(0.00200)
Previous PCI	-0.00985***	-0.00984***	-0.00985***	-0.0124***	-0.0123***	-0.0124***
	(0.00162)	(0.00163)	(0.00162)	(0.00197)	(0.00198)	(0.00198)
Electrophysiology study	-0.0185***	-0.0170***	-0.0176***	-0.0267***	-0.0244***	-0.0249***
	(0.00297)	(0.00286)	(0.00289)	(0.00411)	(0.00392)	(0.00399)
VT indication (ES study)	-0.00531	-0.00446	-0.00481	-0.00742	-0.00606	-0.00654
	(0.00512)	(0.00514)	(0.00510)	(0.00655)	(0.00656)	(0.00665)
Female	-0.0070***	-0.0070***	-0.0071***	-0.0165***	-0.0164***	-0.0164***
	(0.00148)	(0.00148)	(0.00148)	(0.00186)	(0.00186)	(0.00186)
Black	0.0350***	0.0348***	0.0342***	0.0557***	0.0554***	0.0546***
	(0.00242)	(0.00242)	(0.00238)	(0.00300)	(0.00296)	(0.00294)
Hsipanic (Medicare)	0.0124**	0.0115**	0.0121**	0.0183***	0.0169***	0.0169***
	(0.00506)	(0.00505)	(0.00505)	(0.00646)	(0.00646)	(0.00646)
Other race	0.0153***	0.0143***	0.0138***	0.0222***	0.0206***	0.0196***
	(0.00419)	(0.00418)	(0.00417)	(0.00533)	(0.00534)	(0.00533)
Hispanic ethnicity (Registry)	0.00873*	0.00751	0.00621	0.00506	0.00311	0.00216
	(0.00496)	(0.00496)	(0.00499)	(0.00631)	(0.00634)	(0.00639)
2007.year	0.00390	0.00556**	0.00515**	0.00490	0.00755**	0.00687**
	(0.00239)	(0.00240)	(0.00239)	(0.00299)	(0.00302)	(0.00300)
2008.year	0.00906***	0.0111***	0.0105***	0.00809**	0.0114***	0.0104***
	(0.00260)	(0.00261)	(0.00260)	(0.00338)	(0.00340)	(0.00337)
2009.year	0.00845***	0.0107***	0.00980***	0.0105***	0.0141***	0.0128***
	(0.00274)	(0.00275)	(0.00276)	(0.00345)	(0.00344)	(0.00345)
2010.year	0.0123***	0.0143***	0.0132***	0.00713**	0.0103***	0.00873**
	(0.00290)	(0.00290)	(0.00291)	(0.00353)	(0.00357)	(0.00355)
2011.year	0.00755**	0.00922***	0.00783**	0.0108***	0.0134***	0.0114***
	(0.00307)	(0.00311)	(0.00314)	(0.00382)	(0.00386)	(0.00386)
2012.year	0.0130***	0.0141***	0.0125***	0.0138***	0.0155***	0.0131***
	(0.00325)	(0.00327)	(0.00328)	(0.00398)	(0.00403)	(0.00399)
2013.year	0.0129***	0.0142***	0.0123***	0.0193***	0.0214***	0.0189***
	(0.00324)	(0.00325)	(0.00327)	(0.00395)	(0.00400)	(0.00397)
Constant	0.0965***	0.125***	0.187***	0.164***	0.209***	0.293***
	(0.00957)	(0.0100)	(0.0135)	(0.0123)	(0.0126)	(0.0177)
Observations	254,237	254,237	253,596	254,237	254,237	253,596
R-squared	0.034	0.035	0.035	0.057	0.057	0.057

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table A2: Mortality (One & Two Years) Random Effects Regression

VARIABLES	(1) death1yr	(2) death1yr	(3) death1yr	(4) death2yr	(5) death2yr	(6) death2yr
HRR-level ICD Rate	0.0717***	0.100***	0.0848***	0.0734***	0.116***	0.0952***
	(0.0220)	(0.0219)	(0.0218)	(0.0280)	(0.0277)	(0.0276)
Ln(volume)		-0.00852***	-0.00846***		-0.0129***	-0.0130***
		(0.00103)	(0.00102)		(0.00131)	(0.00130)
HRR-level Rx Rate			-0.0930***			-0.126***
			(0.0142)			(0.0183)
Ejection Fraction (EF) <20%	-0.00358***	-0.00358***	-0.00359***	-0.00464***	-0.00465***	-0.00467***
	(0.000341)	(0.000341)	(0.000341)	(0.000427)	(0.000427)	(0.000428)
EF 20-25%	-0.00459***	-0.00459***	-0.00457***	-0.00550***	-0.00549***	-0.00547***
	(0.000394)	(0.000393)	(0.000394)	(0.000493)	(0.000493)	(0.000493)
EF 25-30%	-0.00225***	-0.00226***	-0.00225***	-0.00388***	-0.00390***	-0.00390***
	(0.000383)	(0.000383)	(0.000383)	(0.000480)	(0.000480)	(0.000480)
EF 30-35%	-0.000778*	-0.000769*	-0.000773*	-0.00109*	-0.00108*	-0.00107*
	(0.000465)	(0.000465)	(0.000465)	(0.000582)	(0.000582)	(0.000583)
EF > 35%	0.00150***	0.00150***	0.00149***	0.00198***	0.00198***	0.00195***
	(0.000294)	(0.000294)	(0.000294)	(0.000368)	(0.000368)	(0.000368)
EF Missing	0.0188***	0.0182***	0.0186***	0.0270***	0.0262***	0.0251***
	(0.00690)	(0.00690)	(0.00693)	(0.00864)	(0.00864)	(0.00867)
NY Heart Assoc. Class II	0.00249	0.00246	0.00194	0.00509	0.00505	0.00445
	(0.00394)	(0.00393)	(0.00394)	(0.00493)	(0.00493)	(0.00493)
NY Heart Assoc. Class III	0.0478***	0.0479***	0.0474***	0.0705***	0.0706***	0.0700***
	(0.00387)	(0.00387)	(0.00387)	(0.00484)	(0.00484)	(0.00485)
NY Heart Assoc. Class IV	0.154***	0.153***	0.153***	0.189***	0.189***	0.188***
	(0.00491)	(0.00491)	(0.00492)	(0.00615)	(0.00615)	(0.00615)
NY Heart Assoc. Class missing	0.0507***	0.0492***	0.0492***	0.0870***	0.0849***	0.0852***
	(0.0117)	(0.0117)	(0.0117)	(0.0147)	(0.0147)	(0.0147)
Age 70-74	0.0150***	0.0151***	0.0152***	0.0276***	0.0277***	0.0277***
	(0.00182)	(0.00182)	(0.00182)	(0.00228)	(0.00228)	(0.00228)
Age 75-79	0.0366***	0.0367***	0.0367***	0.0647***	0.0648***	0.0649***
	(0.00183)	(0.00183)	(0.00183)	(0.00229)	(0.00229)	(0.00230)
Age 80-84	0.0625***	0.0626***	0.0628***	0.108***	0.109***	0.109***
	(0.00201)	(0.00200)	(0.00201)	(0.00251)	(0.00251)	(0.00251)
Age 85-89	0.100***	0.100***	0.100***	0.173***	0.173***	0.173***
	(0.00288)	(0.00287)	(0.00288)	(0.00360)	(0.00360)	(0.00360)
Age 90+	0.179***	0.179***	0.179***	0.270***	0.270***	0.271***
	(0.00781)	(0.00781)	(0.00782)	(0.00978)	(0.00978)	(0.00979)
Previous cardiac arrest	0.0563***	0.0559***	0.0558***	0.0557***	0.0550***	0.0552***
	(0.00454)	(0.00454)	(0.00455)	(0.00569)	(0.00569)	(0.00569)
Family history sudden arrest	-0.0115***	-0.0116***	-0.0115***	-0.0183***	-0.0184***	-0.0183***
	(0.00378)	(0.00378)	(0.00378)	(0.00473)	(0.00473)	(0.00473)
Ventricular tacchycardia	0.0446***	0.0446***	0.0444***	0.0569***	0.0569***	0.0567***
	(0.00160)	(0.00160)	(0.00160)	(0.00201)	(0.00201)	(0.00201)
Non-ischemic dilated cardiomyopathy	-0.0202***	-0.0199***	-0.0198***	-0.0299***	-0.0295***	-0.0293***

	(0.00231)	(0.00231)	(0.00231)	(0.00289)	(0.00289)	(0.00290)
Ischemic heart disease	0.0173***	0.0176***	0.0176***	0.0270***	0.0275***	0.0275***
	(0.00246)	(0.00246)	(0.00247)	(0.00308)	(0.00308)	(0.00309)
Previous myocardial infarction	0.00855***	0.00862***	0.00866***	0.0130***	0.0131***	0.0133***
	(0.00165)	(0.00165)	(0.00165)	(0.00207)	(0.00206)	(0.00207)
Previous CABG	0.00883***	0.00885***	0.00868***	0.0204***	0.0205***	0.0203***
	(0.00156)	(0.00156)	(0.00156)	(0.00195)	(0.00195)	(0.00195)
Previous PCI	-0.0100***	-0.00999***	-0.0100***	-0.0123***	-0.0123***	-0.0124***
	(0.00153)	(0.00153)	(0.00153)	(0.00191)	(0.00191)	(0.00191)
Electrophysiology study	-0.0189***	-0.0180***	-0.0183***	-0.0262***	-0.0250***	-0.0251***
	(0.00277)	(0.00277)	(0.00277)	(0.00347)	(0.00346)	(0.00347)
VT indication (ES study)	-0.00436	-0.00401	-0.00400	-0.00620	-0.00563	-0.00563
	(0.00522)	(0.00522)	(0.00522)	(0.00654)	(0.00653)	(0.00654)
Female	-0.00718***	-0.00716***	-0.00716***	-0.0166***	-0.0166***	-0.0165***
	(0.00147)	(0.00147)	(0.00147)	(0.00184)	(0.00184)	(0.00184)
Black	0.0301***	0.0300***	0.0299***	0.0494***	0.0493***	0.0493***
	(0.00225)	(0.00225)	(0.00225)	(0.00283)	(0.00282)	(0.00282)
Hsipanic (Medicare)	0.00903*	0.00867*	0.00957*	0.0138**	0.0133**	0.0136**
	(0.00505)	(0.00505)	(0.00506)	(0.00633)	(0.00632)	(0.00634)
Other race	0.0118***	0.0113***	0.0112***	0.0175***	0.0168***	0.0163***
	(0.00417)	(0.00417)	(0.00417)	(0.00523)	(0.00522)	(0.00522)
Hispanic ethnicity (Registry)	0.00483	0.00443	0.00346	0.000432	-0.000123	-0.000673
	(0.00514)	(0.00514)	(0.00515)	(0.00644)	(0.00644)	(0.00645)
2007.year	0.00387	0.00558**	0.00530**	0.00473	0.00733**	0.00681**
	(0.00256)	(0.00256)	(0.00256)	(0.00316)	(0.00317)	(0.00317)
2008.year	0.00885***	0.0109***	0.0105***	0.00760**	0.0107***	0.0101***
	(0.00265)	(0.00266)	(0.00266)	(0.00328)	(0.00329)	(0.00329)
2009.year	0.00816***	0.0105***	0.00984***	0.00997***	0.0135***	0.0126***
	(0.00266)	(0.00267)	(0.00267)	(0.00330)	(0.00331)	(0.00331)
2010.year	0.0115***	0.0136***	0.0129***	0.00563	0.00874**	0.00768**
	(0.00279)	(0.00279)	(0.00279)	(0.00346)	(0.00346)	(0.00347)
2011.year	0.00677**	0.00829***	0.00746**	0.00915**	0.0115***	0.0102***
	(0.00306)	(0.00305)	(0.00306)	(0.00382)	(0.00380)	(0.00381)
2012.year	0.0122***	0.0132***	0.0122***	0.0121***	0.0137***	0.0120***
	(0.00316)	(0.00315)	(0.00315)	(0.00395)	(0.00392)	(0.00393)
2013.year	0.0122***	0.0134***	0.0122***	0.0176***	0.0195***	0.0178***
	(0.00320)	(0.00319)	(0.00319)	(0.00400)	(0.00397)	(0.00398)
Constant	0.102***	0.130***	0.191***	0.175***	0.216***	0.300***
	(0.00923)	(0.00977)	(0.0135)	(0.0116)	(0.0123)	(0.0171)
Observations	254,237	254,237	253,596	254,237	254,237	253,596
Groups	1,549	1,549	1,543	1,549	1,549	1,543

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table A.3: Mortality (One & Two Year) OLS Fixed Effects Regression

VARIABLES	(1) death1yr	(2) death1yr	(3) death1yr	(4) death2yr	(5) death2yr	(6) death2yr
HRR-level ICD Rate	0.0582*	0.0723**	0.0698**	0.0332	0.0554	0.0531
	(0.0348)	(0.0352)	(0.0353)	(0.0440)	(0.0442)	(0.0445)
Ln(volume)		-0.00567***	-0.00586***		-0.00888***	-0.00925***
		(0.00220)	(0.00222)		(0.00274)	(0.00277)
HRR-level Rx Rate			-0.0713			-0.0986
			(0.0781)			(0.0987)
Ejection Fraction (EF) <20%	-0.00367***	-0.00367***	-0.00369***	-0.00480***	-0.00480***	-0.00483***
	(0.000380)	(0.000381)	(0.000381)	(0.000471)	(0.000471)	(0.000472)
EF 20-25%	-0.00454***	-0.00454***	-0.00452***	-0.00544***	-0.00544***	-0.00541***
	(0.000408)	(0.000408)	(0.000409)	(0.000501)	(0.000501)	(0.000502)
EF 25-30%	-0.00225***	-0.00225***	-0.00225***	-0.00387***	-0.00386***	-0.00387***
	(0.000356)	(0.000356)	(0.000357)	(0.000451)	(0.000451)	(0.000452)
EF 30-35%	-0.000749*	-0.000750*	-0.000757*	-0.00104*	-0.00104*	-0.00102*
	(0.000413)	(0.000412)	(0.000413)	(0.000549)	(0.000549)	(0.000550)
EF > 35%	0.00151***	0.00152***	0.00150***	0.00198***	0.00199***	0.00195***
	(0.000316)	(0.000316)	(0.000316)	(0.000390)	(0.000390)	(0.000390)
EF Missing	0.0158**	0.0158**	0.0165**	0.0243**	0.0242**	0.0238**
	(0.00768)	(0.00768)	(0.00774)	(0.0101)	(0.0101)	(0.0101)
NY Heart Assoc. Class II	0.00226	0.00230	0.00185	0.00497	0.00504	0.00452
	(0.00348)	(0.00348)	(0.00348)	(0.00495)	(0.00495)	(0.00496)
NY Heart Assoc. Class III	0.0478***	0.0479***	0.0476***	0.0708***	0.0709***	0.0706***
	(0.00350)	(0.00350)	(0.00351)	(0.00494)	(0.00494)	(0.00495)
NY Heart Assoc. Class IV	0.154***	0.154***	0.153***	0.189***	0.189***	0.189***
	(0.00585)	(0.00585)	(0.00586)	(0.00731)	(0.00731)	(0.00733)
NY Heart Assoc. Class missing	0.0448***	0.0447***	0.0449***	0.0825***	0.0823***	0.0828***
	(0.0111)	(0.0111)	(0.0111)	(0.0144)	(0.0144)	(0.0144)
Age 70-74	0.0151***	0.0151***	0.0152***	0.0277***	0.0277***	0.0276***
	(0.00166)	(0.00166)	(0.00166)	(0.00206)	(0.00206)	(0.00206)
Age 75-79	0.0365***	0.0365***	0.0366***	0.0647***	0.0647***	0.0647***
	(0.00182)	(0.00182)	(0.00183)	(0.00228)	(0.00228)	(0.00229)
Age 80-84	0.0617***	0.0618***	0.0620***	0.108***	0.108***	0.108***
	(0.00214)	(0.00214)	(0.00214)	(0.00257)	(0.00257)	(0.00258)
Age 85-89	0.0978***	0.0978***	0.0979***	0.170***	0.170***	0.170***
	(0.00330)	(0.00330)	(0.00330)	(0.00411)	(0.00411)	(0.00412)
Age 90+	0.173***	0.173***	0.174***	0.265***	0.265***	0.266***
	(0.0107)	(0.0107)	(0.0107)	(0.0125)	(0.0125)	(0.0125)
Previous cardiac arrest	0.0544***	0.0543***	0.0541***	0.0532***	0.0531***	0.0531***
	(0.00549)	(0.00549)	(0.00549)	(0.00631)	(0.00631)	(0.00632)
Family history sudden arrest	-0.0114***	-0.0115***	-0.0115***	-0.0176***	-0.0177***	-0.0177***
	(0.00362)	(0.00362)	(0.00362)	(0.00467)	(0.00467)	(0.00467)
Ventricular tacchycardia	0.0442***	0.0442***	0.0441***	0.0565***	0.0565***	0.0563***
	(0.00191)	(0.00191)	(0.00191)	(0.00228)	(0.00228)	(0.00228)

Non-ischemic dilated cardiomyopathy	-0.0186***	-0.0186***	-0.0186***	-0.0281***	-0.0280***	-0.0279***
	(0.00243)	(0.00243)	(0.00243)	(0.00315)	(0.00315)	(0.00315)
Ischemic heart disease	0.0185***	0.0185***	0.0185***	0.0285***	0.0285***	0.0286***
	(0.00264)	(0.00264)	(0.00264)	(0.00330)	(0.00330)	(0.00330)
Previous myocardial infarction	0.00906***	0.00908***	0.00916***	0.0132***	0.0132***	0.0134***
	(0.00172)	(0.00172)	(0.00172)	(0.00220)	(0.00220)	(0.00220)
Previous CABG	0.00886***	0.00884***	0.00874***	0.0202***	0.0202***	0.0202***
	(0.00165)	(0.00165)	(0.00165)	(0.00201)	(0.00201)	(0.00201)
Previous PCI	-0.00996***	-0.00997***	-0.0101***	-0.0122***	-0.0122***	-0.0123***
	(0.00163)	(0.00163)	(0.00163)	(0.00200)	(0.00200)	(0.00200)
Electrophysiology study	-0.0193***	-0.0193***	-0.0193***	-0.0259***	-0.0259***	-0.0257***
	(0.00282)	(0.00282)	(0.00282)	(0.00381)	(0.00381)	(0.00382)
VT indication (ES study)	-0.00444	-0.00427	-0.00400	-0.00726	-0.00699	-0.00676
	(0.00514)	(0.00514)	(0.00514)	(0.00659)	(0.00661)	(0.00662)
Female	-0.00726***	-0.00724***	-0.00722***	-0.0169***	-0.0168***	-0.0168***
	(0.00147)	(0.00147)	(0.00147)	(0.00186)	(0.00186)	(0.00186)
Black	0.0257***	0.0256***	0.0258***	0.0444***	0.0444***	0.0448***
	(0.00231)	(0.00231)	(0.00232)	(0.00292)	(0.00293)	(0.00293)
Hsipanic (Medicare)	0.00468	0.00470	0.00573	0.00944	0.00946	0.00986
	(0.00509)	(0.00509)	(0.00510)	(0.00646)	(0.00646)	(0.00648)
Other race	0.00767*	0.00763*	0.00771*	0.0132**	0.0132**	0.0130**
	(0.00424)	(0.00424)	(0.00425)	(0.00551)	(0.00551)	(0.00551)
Hispanic ethnicity (Registry)	0.000784	0.000786	-0.000117	-0.00384	-0.00384	-0.00418
	(0.00498)	(0.00498)	(0.00499)	(0.00632)	(0.00631)	(0.00632)
2007.year	0.00375	0.00480*	0.00485*	0.00439	0.00603*	0.00602*
	(0.00247)	(0.00250)	(0.00250)	(0.00305)	(0.00311)	(0.00310)
2008.year	0.00852***	0.00974***	0.00975***	0.00666*	0.00858**	0.00859**
	(0.00265)	(0.00270)	(0.00271)	(0.00345)	(0.00352)	(0.00353)
2009.year	0.00710**	0.00852***	0.00836***	0.00838**	0.0106***	0.0104***
	(0.00277)	(0.00283)	(0.00284)	(0.00351)	(0.00358)	(0.00359)
2010.year	0.0100***	0.0113***	0.0112***	0.00284	0.00478	0.00463
	(0.00304)	(0.00309)	(0.00309)	(0.00375)	(0.00378)	(0.00378)
2011.year	0.00493	0.00575*	0.00587*	0.00519	0.00647	0.00666
	(0.00344)	(0.00346)	(0.00346)	(0.00444)	(0.00445)	(0.00446)
2012.year	0.0104***	0.0110***	0.0107***	0.00787*	0.00871*	0.00834*
	(0.00375)	(0.00376)	(0.00376)	(0.00473)	(0.00472)	(0.00473)
2013.year	0.0104***	0.0110***	0.0108***	0.0132***	0.0142***	0.0141***
	(0.00378)	(0.00378)	(0.00378)	(0.00473)	(0.00473)	(0.00474)
Constant	0.105***	0.126***	0.172***	0.182***	0.215***	0.279***
	(0.0111)	(0.0136)	(0.0509)	(0.0142)	(0.0175)	(0.0646)
Observations	254,237	254,237	253,596	254,237	254,237	253,596
R-squared	0.045	0.045	0.045	0.045	0.045	0.045

Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A.4: Risk Adjustment Regressions for ICD Rates

A.4.1: Year = 2002

N = 4,964,561 (20% sample).
R-squared = 0.01x

Variable	DF	Parameter Estimate	Standard Error	t-Value
Age 65-69		0.00190	0.00005412	35.08
Age 70-74		0.00196	0.00005403	36.24
Age 75-79		0.00183	0.00005491	33.32
Age 80-84		0.00124	0.00005852	21.17
Female		-0.00175	0.00004916	35.55
Black		0.00010194	0.00011105	0.40
White		0.00028484	0.00008870	3.21
ZIP Income	1	0.00000161	0.00009434	0.02
ZIP Poverty rate		-0.00072561	0.00042969	-1.69
Dual Eligible (yes=1)		-0.00074099	0.00005040	-14.70
County Smoking Rate		-0.00002326	0.00000668	-3.48
County Diabetic Rate		-0.00003132	0.00002310	-1.36
County Obesity Rate		0.00000141	0.00000765	0.18
County Drinking Rate		0.00000285	0.00001061	0.27
Diagnosed CHF (yes=1)		0.00684	0.00004263	160.52

Reference groups: Male, Age 85+, Other race. HRR fixed effects included.

A.4.2: Year = 2006

N = 24,606,213

R-squared = 0.014

Variable	Parameter Estimate	Standard Error	t-Value
Age 65-69	0.00336	0.00003208	104.79
Age 70-74	0.00352	0.00003263	107.77
Age 75-79	0.00342	0.00003307	103.44
Age 80-84	0.00239	0.00003468	68.94
Female	-0.00264	0.00002731	-96.80
Black	-0.00004876	0.00006464	-0.75
White	0.00010578	0.00005058	2.09
ZIP Income	-0.00007070	0.00005687	-1.24
ZIP Poverty	-0.00058706	0.00026231	-2.24
Dual-Eligible (yes=1)	-0.00114	0.00003058	-37.22
County Smoking	-0.00001244	0.00000384	-3.24
County Diabetes	-0.00005567	0.00001212	-4.59
County Obesity	0.00001787	0.00000427	4.19
County Drinking	0.00002902	0.00000707	4.11
Diagnosed CHF (yes=1)	0.01329	0.00002588	513.57

Reference groups: Male, Age 85+, Other race. HRR fixed effects included.

A.4.3. Year = 2013

N = 24,236,818
R-squares = 0.011

Variable	Parameter Estimate	Standard Error	t-Value
Age 65-69	0.00231	0.00002522	91.75
Age 70-74	0.00243	0.00002620	92.73
Age 75-79	0.00236	0.00002755	85.66
Age 80-84	0.00180	0.00002903	62.13
Female	-0.00147	0.00002258	-64.96
Black	0.00000132	0.00004891	0.03
White	-0.00000656	0.00003605	-0.18
ZIP Income	0.00004656	0.00004611	1.01
ZIP Poverty rate	0.00017423	0.00021617	0.81
Dual-eligible (Yes=1)	-0.00065135	0.00002520	-25.85
County Smoking	0.00000428	0.00000341	1.25
County Diabetes	-0.00003945	0.00000854	-4.62
County Obesity	0.00001063	0.00000326	3.26
County Drinking	0.00000949	0.00000529	1.79
CHF Diagnosis (yes=1)	0.01024	0.00002266	451.75