Patient Heterogeneity and Treatment Choices: Evidence from the Dialysis Industry*

Paul J. Eliason†, Benjamin Heebsh,‡ Riley J. League§, Ryan C. McDevitt,¶ James W. Roberts‖

June 22, 2026

We study the use and effectiveness of epoetin alfa (EPO), historically one of Medicare's largest drug expenses, by estimating a model of providers' treatment decisions that incorporates the heterogeneous effect of EPO on patients. Increasing doses by 1% reduces anemia by 0.1%, on average, but increases heart attacks by 0.3% and mortality by 0.4%, with patients exhibiting substantial heterogeneity along these dimensions. By estimating the distribution of patient-specific treatment effects of EPO and how providers optimally respond to them, our analysis shows that patients' well-being would decline considerably if providers followed guidelines that disregard these heterogeneous responses.

JEL Codes: D22, I11, I18, L20

*We thank nephrologists Matthew Ellis, Ruediger Lehrich, John Middleton, and Myles Wolf for providing valuable industry and medical insights. We gratefully acknowledge research support provided by the Social Sciences Research Institute, the National Science Foundation (SES 1850736 and 1851615), and the National Institute on Aging (T32-AG000186). We thank Jeffrey Hill, Alexander Marsh, Gabor Palinko, Lily Liu, Zilan Yang, Sungwoo Cho, Hanmeng Wang, and Wanrong Deng for their research assistance. The data reported here have been supplied by the United States Renal Data System (USRDS). The interpretation and reporting of these data are the responsibility of the authors and in no way should be seen as an official policy or interpretation of the U.S. government.

†David Eccles School of Business, University of Utah and NBER, paul.eliason@eccles.utah.edu

‡Federal Trade Commission, bheebsh@ftc.gov. The views expressed in this article are those of the authors and do not represent the views of the Federal Trade Commission (FTC) or any of its commissioners.

§Gies College of Business, University of Illinois Urbana-Champaign and NBER, league@illinois.edu

¶Washington University and NBER, mcdevitt@wustl.edu

‖Department of Economics, Duke University and NBER, j.roberts@duke.edu

1. Introduction

We study the use and effectiveness of epoetin alfa (EPO), a drug commonly prescribed to treat anemia among dialysis patients that was Medicare's largest Part B drug expense in the early 2000s (U.S. Government Accountability Office, 2012; Thamer et al., 2007). We estimate a structural model in which providers face potentially conflicting objectives as dosing decisions depend on their own financial returns as well as the health of their patients, captured through a health production function that incorporates both observed and unobserved components of patients' health and allows EPO to have heterogeneous effects across patients. Modeling the health effects of EPO in this way quantifies the tradeoffs faced by providers and allows us to demonstrate how they depend on facility attributes such as corporate ownership and the costs of acquiring the drug, two important features of the dialysis industry that directly affect Medicare's payment policies. Even though misaligned financial incentives distort treatment decisions, we find that allowing providers to tailor EPO doses to individual patients' needs results in better outcomes.

Identifying the parameters of the health production function is complicated by the fact that providers base their treatment decisions in part on patients' underlying health, so any correlation between doses and outcomes may be confounded by patients' unobserved characteristics. Reflecting this possibility, OLS regressions of hemoglobin and blood transfusions on patients' EPO doses produce spurious negative and positive correlations, respectively, despite evidence from clinical trials that show the drug has the opposite effect on these measures of anemia (Eschbach et al., 1989). Moreover, evidence from clinical trials cannot be used to calibrate our health production function, as these trials provide little information about the heterogeneous effect of EPO and use patient populations and dosing regimens that do not fully reflect how the drug is used in practice.

To overcome the empirical challenges associated with patients' unobserved health conditions and coincidental changes in their level of care, we use two novel sources of exogenous variation in providers' treatment decisions to identify the marginal effect of EPO on health outcomes. First, Medicare implemented a major payment reform for EPO during our sample period. Prior to 2011, Medicare paid dialysis facilities a fee-for-service reimbursement for the amount of EPO administered during treatment, providing an incentive for providers to use excessive doses and raising concerns about an increased risk of mortality and cardiovascular events (Besarab et al., 1998; Singh et al., 2006). In response to the es-

calating costs of EPO and its potentially harmful effect on patients, Medicare switched to a prospective payment system (PPS) in 2011 to encourage providers to use the drug more judiciously, paying for both EPO and dialysis together with a single, predetermined payment.

Second, dialysis patients living at higher elevations have less severe anemia at baseline and therefore require less EPO to manage their condition (Winkelmayer et al., 2009; Brookhart et al., 2011). During the fee-for-service era, this physiological distinction made patients at higher elevations less profitable for providers, as they received smaller doses of EPO to keep their blood levels within clinical guidelines. After Medicare adopted the PPS, providers' financial incentives flipped, with patients at lower elevations becoming less lucrative for providers who no longer received incremental payments corresponding to these patients' larger doses. As a result, the uniformly applied payment reform ultimately had different financial implications for facilities at different elevations and resulted in different changes in patients' doses of EPO.

Using the interaction between the payment reform and a facility's elevation as an instrumental variable, we estimate the heterogeneous effect of EPO on patients' health. We find that, in equilibrium, patients exhibit substantial heterogeneity in their response to the drug, with patients at the 75th percentile of effectiveness responding more than six times as strongly as those at the 25th. Providers' treatment decisions also come with important tradeoffs: a 1% increase in EPO reduces patients' anemia by 0.1%, on average, but increases the rate of cardiac hospitalization by 0.3% and mortality by 0.4%. How providers weigh these tradeoffs against their own financial incentives ultimately determines the amount of EPO they administer. Because these causal effects vary by patient, we can recover the preference weights providers place on each outcome from variation in doses — and how those doses changed following the payment reform — across patients who differ in their likelihood of benefiting from or being harmed by EPO.

The resulting estimates form the basis of our counterfactual analysis that explores how patient heterogeneity, financial incentives, and regulatory constraints together influence treatment decisions and patients' well-being. In our analysis, we show that even though financial incentives distort treatment decisions, allowing providers to tailor their doses to each patient's individual needs improves outcomes, which has important policy implications given EPO's historic overuse by for-profit chains (Eliason et al., 2020). Before Medicare's move to a PPS, EPO had accounted for as much as 25% of revenue at DaVita, one of the nation's two largest chains, and up to 40% of its profits (DaVita, 2005), with lawsuits at the

time alleging the company developed a software application called Snappy to automatically increase doses so patients' blood levels hit the upper bound of clinical guidelines and maximize fee-for-service reimbursements (Mueller, 2023).

From our estimates, we find that a policy that requires all patients to receive enough EPO to achieve the same minimum hemoglobin level — i.e., the level recommended by clinical guidelines — would cause health to decline. We further demonstrate the value of allowing providers to tailor their doses to patients' specific needs by quantifying the impact of using a uniform dose for all patients, set at the average health-maximizing dose, finding that such a constraint would harm health by an amount valued at nearly $800 per patient-month. Given the challenges of contracting care for patients with diverse needs and responses, this illustrates the importance of allowing for provider discretion in treatment decisions, even when misaligned financial incentives might spur policymakers to impose guidelines to restrain costs and protect patients. Medicare's Quality Incentive Program is one such case: at times it has penalized providers when their patients' hemoglobin levels fall outside clinical guidelines regardless of whether a given patient actually benefits from being within the recommended range (Bertuzzi et al., 2023).

Our model's estimates also allow us to evaluate how provider heterogeneity influences patient care. With large chains such as DaVita and Fresenius now operating nearly 80% of facilities nationwide (Xia et al., 2025), two factors in particular distinguish their treatment decisions from those of independent facilities. First, we find that chains put less emphasis on patients' well-being when administering EPO: patients' health would improve enough to close 59% of the gap between observed outcomes and the theoretical upper bound if chains had the same level of altruism as independent facilities do. Conversely, chains' lower acquisition costs for EPO spur them to use larger doses, benefiting patients who might otherwise be undertreated due to the financial incentives of the PPS. By highlighting the interaction between Medicare's payment policies and the corporate ownership of facilities, our findings provide novel insights for policymakers seeking to implement similar reforms on chains throughout the broader health care system.

Our study contributes to the large literature aimed at understanding the behavior of health care providers, including their treatment decisions, quality of care, and associated costs (McGuire, 2000). One strand of this work seeks to explain the wide variation in treatment decisions observed in practice (Chandra et al., 2012). Because treatment variation is often unrelated to providers' quality of care, some

point to it as evidence of the health care system's inefficiency, although a more recent literature has shown that a meaningful portion of the variation stems from differences in providers' productivity (Chan et al., 2022; Silver, 2021; Chandra and Staiger, 2020; Currie and MacLeod, 2017; Abaluck et al., 2016). Advancing this literature, we emphasize the role of patients' heterogeneous response to treatments and how it relates to providers' and regulators' various objectives (Dafny, 2005; Clemens and Gottlieb, 2014; Einav et al., 2018; Eliason et al., 2018).

Using observed provider and patient behavior to study the effectiveness of EPO also contributes to a growing literature exploring the limitations of clinical trials for assessing interventions in the field. To highlight one such limitation, trial participants may differ from the general population and fail to represent important subgroups (e.g., Murthy et al., 2004; Chassang and Feng, 2022; Abaluck et al., 2025), with a failure to distinguish treatment effects for these groups potentially limiting our understanding of the intervention and how it will ultimately affect spending and outcomes (Green et al., 2022). Gaps in representation also contribute to disparities in care, including slower uptake of newly developed drugs among groups with less representation in trials (Alsan et al., 2022) and limited guidance for how to adapt care for underrepresented populations (Cuddy and Currie, 2024). In addition, design features of controlled trials, such as site selection (Allcott, 2015) and adherence protocols (Al-Ubaydli et al., 2017a), may appear to enhance the effectiveness of an intervention and abstract away from institutional or equilibrium factors, such as financial incentives, that arise only outside of the lab or when the intervention is deployed at scale (Oostrom, 2022; Bold et al., 2018; Al-Ubaydli et al., 2017b). For dialysis in particular, clinical trials of EPO have demonstrated its ability to reduce anemia (Eschbach et al., 1987) as well as raising concerns about adverse effects for certain groups of patients (Besarab et al., 1998; Singh et al., 2006). Our paper builds on this work by using data from outside a controlled environment, showing evidence of wide heterogeneity in treatment effects and highlighting the role of providers' incentives in shaping equilibrium outcomes. We also contribute to the understanding of why some patients exhibit a diminished responsiveness to EPO, confirming the hypothesis that certain comorbid conditions, such as hypertension, can significantly reduce the effectiveness of the drug (van der Putten et al., 2008).

Finally, our paper contributes to a recent literature focused on the economics of the dialysis industry (e.g., Eliason et al., 2020; Dai, 2014; Cutler et al., 2017; Dai and Tang, 2015; Grieco and McDevitt, 2017; Eliason, 2021; Wilson, 2016a,b), including research on the switch to a PPS (Chertow et al., 2016; Hirth

et al., 2014). Of particular relevance, Gaynor et al. (2023) use a structural model to study how dialysis providers balance patients' health with the financial incentives related to administering EPO. Using data from before the payment reform, they find that fee-for-service payments resulted in an excessive use of EPO, with doses falling by 10–32% under an optimal contract. We complement their work by using the payment reform and an instrumental variables strategy to identify the health production function and estimate a model of provider treatment decisions, using it to explore the role of patient heterogeneity in determining EPO doses. We also extend their results by measuring the impact of EPO on health outcomes such as hospitalization and mortality.

Our paper proceeds with Section 2, which discusses essential details of the U.S. dialysis industry. Section 3 describes our data and provides descriptive statistics. Section 4 lays out a model of providers' equilibrium dosing decisions. Section 5 discusses our estimation strategy. Section 6 presents our model's estimation results, including an IV estimation of EPO's effect on patients' outcomes. Section 7 reports estimates from various counterfactual simulations. Section 8 concludes. The appendix contains supplementary material referenced throughout the paper.

2. Institutional Details of Dialysis

2.1. Kidney Failure, Anemia, and Dialysis

Kidneys filter wastes and toxins from the blood and produce erythropoietin, a hormone that stimulates red blood cell production. For patients with end-stage renal disease (ESRD), the kidneys no longer adequately perform these functions. To survive, those with ESRD must either receive a kidney transplant or go on dialysis, a medical treatment that clears wastes and toxins from the blood.

The most common form of dialysis, hemodialysis, uses a machine to artificially clean blood outside the body, either at the patient's home or at a medical facility. Because over 90% of dialysis patients in the U.S. use in-center hemodialysis, we focus on that modality for our analysis.[1] In-center hemodialysis patients primarily receive treatment at one of the more than 7,000 dedicated dialysis facilities across the country, where they typically go three times per week for treatment that lasts three to four hours each visit. These facilities are run by a mix of for-profit and non-profit firms, with the two largest for-profit

[1]Please see Wang et al. (2018) for a discussion of the trends in dialysis modalities.

chains, DaVita and Fresenius, controlling nearly 80% of facilities and earning 90% of the industry's revenue (Xia et al., 2025; Baker, 2019). Independent facilities comprise most of the remainder.

In addition to needing wastes and toxins cleared from their blood, ESRD patients often require other treatments to address complications associated with kidney failure. For instance, nearly all ESRD patients suffer from anemia (Babitt and Lin, 2012). In healthy kidneys, specialized cells detect hypoxia (i.e., low oxygen levels) in the bloodstream and release erythropoietin, a protein that stimulates bone marrow to make red blood cells, the effectiveness of which depends on the availability of iron and other substrates in the bloodstream (Nakhoul and Simon, 2016).[2] In patients with ESRD, however, kidneys usually fail to perform this function.

Among dialysis patients, anemia is typically managed using drugs administered by the dialysis facility during a session. Chief among them is recombinant human erythropoietin, or epoetin alfa, a biologic commonly known as EPO that supplements naturally occurring erythropoietin. Manufactured by Amgen under the brand name EPOGEN, EPO was approved by the Food and Drug Administration (FDA) in 1989 to treat anemia in dialysis patients (Kalantar-Zadeh, 2017) and became the standard of care for the condition, with those treated with EPO requiring fewer blood transfusions and reporting improved appetite, activity level, and sense of well-being (Eschbach et al., 1989; Valderrabano, 2000). By 2005, 99% of dialysis patients regularly received EPO, and in some years it was Medicare's largest Part B drug expense (Thamer et al., 2007).

By the mid-2000s, randomized controlled trials began to show that EPO may be harmful under certain conditions. Besarab et al. (1998), for instance, found that ESRD patients with congestive heart failure had a higher probability of myocardial infarction and death when treated with EPO to achieve normal or high hematocrit levels, while Singh et al. (2006) found evidence of an increased risk of cardiovascular events and death among pre-dialysis chronic kidney disease patients when they received enough EPO to achieve normal or high hematocrit levels. Although these trials focused on specific patient populations, they raised concerns about the use of EPO more generally, and in March 2007 the FDA issued a public health advisory for the drug, mandating a black box warning and advising physicians to adjust doses to target hemoglobin (HGB) levels between 10 and 12 g/dL (Thamer et al., 2013). Over this period, observational studies suggested similar adverse effects (Zhang et al., 2004;

[2]In addition to iron, essential substrates of red blood cells include vitamin B12, B6, folate, and amino acids, although shortages of these other substrates are relatively rare, even among dialysis patients (Badura et al., 2024).

Bradbury et al., 2009; Brookhart et al., 2010), although they did not use a research design that produced causal estimates and providers did not change their doses much in response (Thamer et al., 2013). In June 2011, the FDA amended the original black box warning to include guidance to dose no higher than what would be necessary for the patient to avoid a blood transfusion.

Notably, not all patients require the same amount of EPO to hit a given HGB target (Kalantar-Zadeh et al., 2009). For instance, patients with higher baseline HGB levels (i.e., the HGB they would have in the absence of any EPO treatment) require less EPO to increase their HGB to meet clinical guidelines. In addition, a host of factors correlate with a patient's responsiveness to EPO, such as comorbid conditions associated with inflammation, which depress the availability of iron and dampen the effectiveness of EPO by limiting the ability of bone marrow to produce red blood cells (Nakhoul and Simon, 2016).

The elevation at which a patient resides provides another source of variation in the amount of EPO needed to achieve a given HGB target. Observational studies have shown that patients residing at higher elevations achieve higher HGB levels, even while receiving lower doses of EPO (Winkelmayer et al., 2009; Brookhart et al., 2011). Two distinct mechanisms potentially explain this correlation. First, dialysis patients at higher elevations tend to have higher baseline levels of HGB because the richness of oxygen in their blood is lower, which activates an increase in both naturally occurring erythropoietin and the amount of iron in the bloodstream, meaning that these patients can achieve their desired HGB target with smaller doses of EPO.[3] Consistent with this mechanism, experimental studies artificially replicating elevation-induced hypoxia using hypobaric chambers have shown that patients diagnosed with chronic kidney failure produce more natural erythropoietin at elevation (Bosman et al., 2002). In other words, naturally occurring erythropoietin raises baseline HGB levels and crowds out the need for supplemental EPO, even if the effectiveness of EPO remains unchanged. Adding further support for this mechanism, we show in Appendix Section E.3 that incident patients at higher elevations without prior anemia treatment have higher baseline HGB levels.

As a second potential mechanism, elevation could make EPO itself more effective. The medical evidence on this possibility is less conclusive. Theoretically, patients' iron availability could naturally increase with elevation, and more iron could make EPO more effective (Nakhoul and Simon, 2016).

[3]See Winkelmayer et al. (2009) and Brookhart et al. (2011) for a more complete discussion of these physiological relationships.

Observational studies of dialysis patients suggest otherwise, however, finding that dialysis patients at different elevations have similar baseline levels of iron availability (Sibbel et al., 2017).[4] In Appendix Section E.3, we show that iron availability also does not vary by elevation among patients initiating dialysis in our sample and provide further supporting evidence that the effect of EPO does not vary with elevation. In short, existing evidence for why patients at higher elevation need less EPO supports the first mechanism, that elevation affects baseline HGB levels, and contradicts the second, that elevation affects EPO's effectiveness.

2.2. Medicare Payment Reform

Since 1972, Medicare has extended full benefits to all patients with ESRD, regardless of age, paying for both dialysis and anemia treatment under Part B. Those enrolled in an employer group health plan when diagnosed with ESRD retain their commercial insurance as the primary payer for 33 months, during which time Medicare acts as a secondary payer (Lin, 2021; League et al., 2022).

From the early 1980s to 2010, Medicare paid providers a composite rate of approximately $128 per dialysis session, an amount intended to cover the labor, capital, and routine lab tests associated with each treatment, in addition to a fee-for-service rate for EPO and other injectable drugs. Initially, Medicare set the payment rate for EPO at $10 per 1000 IUs and then updated the rate in 2005 based on the average sales price plus a 6% markup, resulting in a slight decline in payments to about $9.50 per 1000 IUs.[5] EPO doses and expenditures increased consistently during the fee-for-service era, with spending on erythropoiesis-stimulating agents such as EPO approaching $2.7 billion in 2007 (Whoriskey, 2012). Concerns that the distortionary incentives from fee-for-service payments resulted in excessive costs for Medicare and harm to patients motivated policymakers to include ESRD payment reform as part of the Medicare Improvements for Patients and Providers Act (MIPPA) in 2008.

MIPPA mandated the bundling of dialysis services and all injectable drugs and biologics used in the treatment of ESRD into a single prospective payment, to take effect in 2011. For the first year

[4]Experimental studies simulating altitude for healthy patients present evidence of increased iron mobilization — that is, iron stores being transferred to the bloodstream for red blood cell production — with higher elevation (Gassmann and Muckenthaler, 2015). The increase in natural EPO at elevation triggers increased red blood cell production, however, and exhausts this additional iron. Sibbel et al. (2017) show that ferritin, a measure of iron storage, and transferrin saturation, a measure of iron available in the bloodstream for red blood cell production, are similar across different elevations. The similarity of equilibrium levels of iron availability across elevations shuts down this potential pathway of enhanced EPO effectiveness.

[5]For more details, please see `https://www.gao.gov/assets/260/253347.pdf`.

of the PPS, Medicare fixed payment for these services at about $230, a level chosen to reduce total federal payments to dialysis providers by 2%.[6] Although targeted primarily at EPO due to the drug's outsize impact on Medicare expenditures and patient outcomes, the original PPS also bundled together 21 other drugs, spanning anemia treatment, access management, and anti-infectives.[7]

To offset the incentives for providers to increase their profits by providing lower-quality care following the payment reform, MIPPA also mandated the development of the ESRD Quality Incentive Program (QIP), which reduces payments to providers who fail to meet certain clinical standards, such as hemoglobin levels and hospitalization rates. Although the specific criteria assessed in the QIP change from year to year, in its inaugural year, 2012, the QIP targeted patients' hemoglobin levels and urea reduction ratio, a measure of dialysis filtration. Under the QIP, Medicare reduces the annual payments to a facility by between 0.5% and 2.0% if, for instance, the HGB levels of too many patients fall outside the regulated standards, with the size of the penalty determined by the extent of the shortfall. We discuss the QIP further in Appendix B, where we provide evidence that it did not meaningfully contribute to the reduction in EPO administered during our sample period, and evaluate providers' strategic responses to it in Bertuzzi et al. (2023).

3. Data, Descriptive Statistics, and Time Trends

The main dataset for our analysis comes from the U.S. Renal Data System (United States Renal Data System, 2019), a national clearinghouse that collects and manages data from a variety of sources relevant to ESRD patients and health care providers. Included in these data are Medicare claims, treatment histories, patient attributes, and annual facility surveys. In addition, CMS Form 2728, known as the Medical Evidence Form, provides data on the health and clinical attributes of patients when they begin dialysis. We also geocode facilities' addresses and extract their elevations using data from the U.S. Geological Survey (U.S. Geological Survey Center for Earth Resources Observation and Science, 2014).

Table 1 presents summary statistics for our variables of interest. We limit our sample to hemodialysis patients between the ages of 18 and 100 who have Medicare as their primary payer and no missing

[6] See Federal Register, Volume 74, Issue 187, (September 29, 2009). Providers had the option to transition into the PPS either immediately in 2011 or gradually over four years starting in 2011. Over 90% opted for the immediate transition. In Appendix A, we demonstrate that our results are robust to using only the set of providers who opted in immediately.

[7] Since then, this list has been expanded to include over 50 products.

observations for the patient and facility characteristics used in our later analysis. These characteristics include demographic variables such as gender and age, comorbidities such as diabetes and cancer, patient behaviors such as smoking and drinking, and facility characteristics such as chain affiliation and elevation. We restrict our analysis to the years 2009–2012, a window centered narrowly around the start of the PPS, leaving us with approximately 10 million patient-month observations. As will be important for our instrumental variable analysis in Section 6, the elevation of facilities varies considerably, with a standard deviation of 924 ft. We present summary statistics by elevation in Appendix C.

Figure 1 shows the evolution of several aspects of the dialysis services included in the prospective payment in 2011, including the quantity and quality of anemia management and dialysis care, before and after the reform and across elevations. Measures of anemia management responded immediately to the payment change: EPO doses were relatively flat going into 2010 but decreased sharply starting midway through the year, moving in concert with the increase in transfusions shown in panel (b), a correlation consistent with EPO being used to increase patients' HGB levels and reduce the need for transfusions. The sharp decline in EPO predates the payment reform in 2011 by a few months and matches Medicare's announcement of the final PPS rule.[8,9]

The dashed lines in Figure 1 show how treatment evolved for patients residing in the first and fifth elevation quintiles. Patients at high elevations received smaller doses of EPO in both the pre- and post-2011 periods, as shown in panel (a), while those at lower elevations experienced a much larger decline following the payment reform. Directly related to the drop in EPO, panel (b) shows a much larger increase in transfusions for patients at lower elevations despite having a rate similar to those at higher elevations prior to 2011.

In contrast to the changes in patients' treatment for anemia, we find little evidence that dialysis care itself changed in response to the payment reform, providing reassurance that we can use the switch to prospective payments to identify the health effects of EPO. As examples, the average number of monthly dialysis sessions per patient in panel (c) remained steady throughout the payment reform, and hospitalizations for septicemia in panel (d), a class of infections that can arise from improper cleaning

[8]We show in Appendix D that our results are robust to changing the treatment period to also include the anticipatory period between the announcement and the implementation of the payment reform. For more details, please see `https://www.gao.gov/assets/gao-13-190r.pdf`.

[9]As discussed in Section 2, two other policy changes occurred during this period, a black box warning and the QIP. In Appendix B, we present evidence that these changes do not explain the decline in EPO doses in Figure 1 — if anything, they would make our estimate of the change in EPO doses attributable to the payment reform conservative.

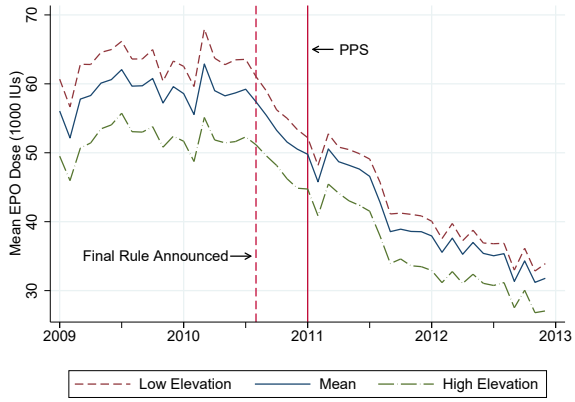
Table 1
Patient Descriptive Statistics

	Mean	Std. Dev.
Patient Characteristics		
Predicted Mortality	0.016	0.011
Age (Years)	63.40	14.57
Months with ESRD	45.08	38.01
Black	0.385	0.487
Male	0.552	0.497
Diabetic	0.540	0.498
Hypertensive	0.906	0.292
Incident Hemoglobin	9.853	1.674
Facility Characteristics		
Facility Elevation (ft)	638.1	923.5
Independent Ownership	0.197	0.397
EPO Use		
EPO Dose (1000 IUs)	48.50	64.11
Receives Any EPO	0.755	0.430
Health Outcomes		
Hemoglobin (g/dL)	11.12	1.22
Transfusions	0.028	0.166
Mortality	0.016	0.124
Hospitalizations		
Any Cause	0.1380	0.3449
Cardiac Event	0.0271	0.1625
Septicemia	0.0094	0.0965
Unique Patients	461,477	
Patient-Months	10,077,289	

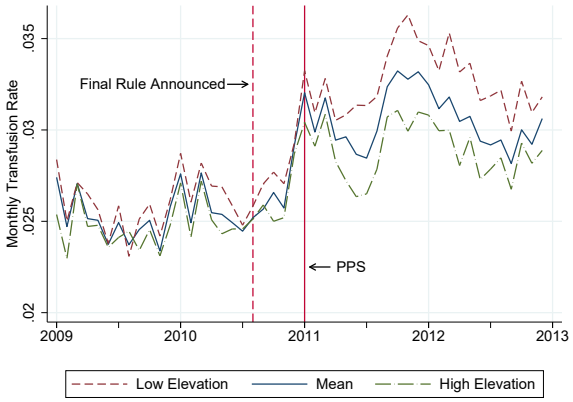
Notes: An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Predicted mortality is the predicted value for each observation using coefficients from a regression of mortality on patient controls and time fixed effects on observations from 2009 and 2010. Time fixed effects are not included in the prediction. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. Transfusion and hospitalization variables are binary indicators. Facility elevation is measured in feet above sea level.

Figure 1
Time Trends in Treatments and Outcomes, by Elevation

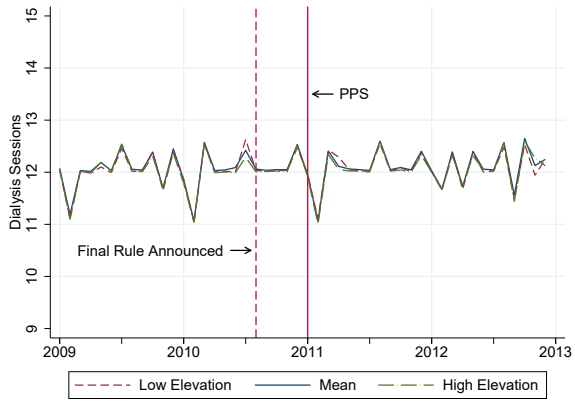
(a) Monthly EPO Dose



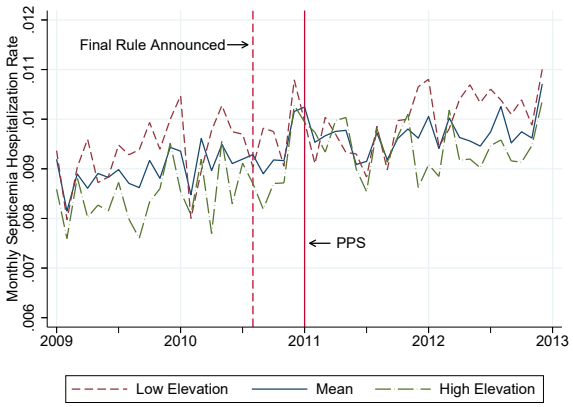
(b) Transfusion Rate



(c) Monthly Dialysis Sessions



(d) Hospitalizations for Septicemia



Notes: An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Transfusion and hospitalization variables are binary indicators. High (low) elevation denotes facility elevation in the fifth (first) quintile. This corresponds to being above 870 (below 73) feet above sea level. The solid vertical line indicates the start of the PPS in January 2011, while the dashed vertical line indicates the announcement of the final rule for the PPS.

of dialysis facilities and reflects low-quality care, did not change.

4. Model of Provider's Treatment Decisions

Our descriptive results from Section 3 suggest the payment reform had a large impact on the amount of EPO administered to dialysis patients and their corresponding measures of anemia management. To better understand how EPO influences patients' health, as well as the tradeoffs providers make between their own financial returns and their patients' well-being, we consider a stylized model of providers' dosing decisions that incorporates the key factors governing these choices and rationalizes their responses to the PPS documented in Section 3.

A dialysis provider must decide how much EPO, *epo*, to administer to a patient with observable characteristics x, unobservable characteristics u, and living at elevation a. Giving the patient *epo* units of EPO influences various components of health, h_c, such as the risk of hospitalization and mortality, given by

$$h_c(epo, x, a, u) = \alpha_c(x)epo + b_c(x, a, u). \tag{1}$$

Here, α_c captures the potentially heterogeneous effect of EPO on outcome c, and b_c is the baseline level of the health component. Following the literature (e.g., Ellis and McGuire, 1986; Gaynor et al., 2023), we assume the patient accepts the amount of EPO recommended by the provider.

In deciding how much EPO to administer, an altruistic provider maximizes their objective function, which takes into account their profit from administering EPO as well as its impact on the patient's health:

$$\begin{aligned} U(epo; p, x, a, u) &= epo * p \;+\; H(epo, x, a, u) \\ &= epo * p \;+\; \sum_c \beta_c h_c(epo, x, a, u) \\ &\quad +\; h_u(epo, x, a, u) \\ &\quad -\; (\beta_{hgb} + \lambda b_{hgb}(x, a, u))(h_{hgb}(epo, x, a, u) - \tau(x, a, u))^2, \end{aligned} \tag{2}$$

where p is the profit margin from a dose of EPO, β_c captures the provider's preferences over the patient's

health for observable component c, excluding HGB, and the final term captures the way that providers consider the impact of EPO on a patient's HGB. In this setup, $H(epo, x, a, u)$ captures all components of patient health and the provider's preferences over it, including unobservable health $h_u(epo, x, a, u)$.

Following Gaynor et al. (2023), we specifically incorporate the impact of EPO on HGB into the provider's dosing decision. Based on the medical literature and industry norms of targeting specific levels of HGB for each patient, we follow Gaynor et al. (2023) and use a quadratic loss function in the difference between total achieved HGB and a target level, τ. We extend the specification in Gaynor et al. (2023) to incorporate the relationship between elevation and HGB by allowing b_{hgb}, the baseline HGB level a patient would have without any EPO, to differ by elevation but do not allow the effect of EPO on HGB to differ by elevation. These assumptions reflect the evidence from the medical literature discussed in Section 2.1, and further supported in Appendix Section E.3, that dialysis patients residing at higher elevations naturally achieve higher baseline levels of HGB, which affects the amount of EPO they need to reach a given target, but patients at higher elevations do not have different levels of iron availability, which would otherwise affect the marginal impact of EPO. To add further flexibility, we allow both the HGB target and the health cost of deviating from this target, $\beta_{hgb} + \lambda b_{hgb}(x, a, u)$, to depend on the patient's baseline HGB level, which is strongly correlated with elevation.

Importantly, we also enrich the health production function in Gaynor et al. (2023) by including other components of a patient's health in addition to HGB. Doing so allows us to model how providers consider the impact of increasing EPO doses on outcomes like cardiac events and mortality.[10]

Finally, in addition to including observable health components, we allow for an unobserved component using the function

$$h_u(epo, x, a, u) = \alpha_u epo + b_u(x, a, u),$$

where EPO benefits health in a way we model as a random effect with standard deviation σ. Note that we normalize the units of unobserved health to be valued by the provider as one dollar of profit.

Our stylized model highlights several important aspects of a provider's dosing decision. Notably,

[10]The RCTs studying the side-effects of EPO are limited in their ability to distinguish whether the risks arise from having high HGB levels or from the EPO itself because they randomized HGB targets and not doses. However, several pieces of evidence suggest that EPO itself comes with inherent risks, including a heightened risk of mortality and cardiac events, even for patients with similar HGB levels. For example, EPO increases blood pressure (Krapf and Hulter, 2009) and activates platelets in the blood (Farag et al., 2012), which could lead to adverse outcomes including heightened risk of stroke, cardiac events, or mortality. We therefore allow providers to value the health impacts arising from HGB separately from side effects.

the first-order condition,

$$g(epo^*; p, x, a, u) = p + \frac{\partial H}{\partial epo^*}(epo^*, x, a, u) = 0, \tag{3}$$

and equilibrium EPO dose equation,

$$epo^* = \frac{p + \sum_c \beta_c \alpha_c(x) + \alpha_u}{2\alpha_{hgb}^2(x)(\beta_{hgb} + \lambda b_{hgb}(x, a, u))} + \frac{\tau(x) - b_{hgb}(x, a, u)}{\alpha_{hgb}(x)}, \tag{4}$$

show that a financially motivated provider will overtreat patients if the profit margin of EPO is positive and undertreat them if it is negative. This relationship mirrors policymakers' concerns both before the payment reform, when they suspected providers administered excessive doses of EPO to generate more revenue, and after, when they feared the reform would result in insufficient doses by making the drug an uncompensated cost for providers.

Second, the model highlights the multidimensional nature of health, whereby providers must trade off the positive effect of EPO on one health outcome against its negative effect on another. In (4), the first term highlights that each non-HGB component of health impacts the dosing, while the second term reflects the distance between the realized HGB level and the target; the existence of other health components means providers shade away from the HGB target.

The model also highlights the role of heterogeneity in patients' responses to EPO. The equilibrium level of EPO changes with respect to its profit margin at the rate

$$\frac{\partial epo^*}{\partial p} = \frac{1}{2\alpha_{hgb}^2(x)(\beta_{hgb} + \lambda b_{hgb}(x, a, u))}, \tag{5}$$

which shows the factors that dictate the extent to which EPO doses adjust in response to a payment change: the adjustment depends on the effectiveness of EPO in changing HGB as well as how providers value patients' health, with more-altruistic providers — or patients for whom hitting the EPO target is more important — responding less strongly.

The data clearly reflect this heterogeneity. Panel (a) of Figure 1, for instance, shows the change in EPO doses differed across elevations following the payment reform. Through the lens of the model, the

influence of elevation on providers' responses to their financial incentives is given by

$$\frac{\partial^2 epo^*}{\partial p \partial a} = -\frac{\lambda \frac{\partial b_{hgb}}{\partial a}(x,a,u)}{2\alpha_{hgb}^2(x)(\beta_{hgb} + \lambda b_{hgb}(x,a,u))^2}, \tag{6}$$

where the smaller change in EPO doses for patients at higher elevations in Figure 1 reveals this value is negative. Because EPO increases HGB and patients' baseline HGB increases with elevation, this implies λ is positive. In other words, the higher a patient's baseline level of HGB, the more costly it is to the patient's health to be farther away from the HGB target, as would be expected given that exceeding the HGB target will be costlier for those with more natural EPO.[11]

In short, the observed differences in how providers respond to payment changes demonstrate that the heterogeneous relationship between health and EPO influences their response. Beyond elevation, other relevant sources of patient heterogeneity could include comorbidities like obesity or heart disease that may similarly impact the effectiveness or risks of EPO and therefore also influence providers' dosing decisions.

5. Identification and Estimation

Estimation proceeds in two stages. First, we recover the production function for each observable component of health $\alpha_c(x)$ using an instrumental variables strategy. Second, we use the heterogeneous effects of EPO identified in the first stage to recover the remaining parameters, which we estimate by GMM. We discuss the identification and estimation of the production function for each observable component of health in Section 5.1 and estimation of the remaining parameters in Section 5.2.

5.1. Estimation of the Health Production Function

A key component of our model is the health production function, which characterizes how EPO affects specific health outcomes. We use a novel instrumental variable strategy to identify the causal

[11] While we do not impose a sign on λ, clinical guidelines and the medical literature suggest we should expect it to be positive, reflecting the asymmetric risks posed by missing the HGB target. Patients with high baseline HGB levels may more easily surpass the HGB targets, leading to greater health risks such as heart failure and myocardial infarction with less benefit. Patients with lower baseline HGB, on the other hand, are more likely to have low HGB, typically facing relatively benign risks related to quality of life and transfusion risks (Kidney Disease: Improving Global Outcomes (KDIGO) Anemia Work Group, 2012).

effect of EPO. To motivate our approach, we first use our instrument to recover the local average causal effect of EPO on the most prominent observable dimensions of health. We then use this same strategy to estimate our production function to allow for a greater degree of heterogeneity. Because estimating these effects does not depend on the functional form assumptions we make when estimating the remainder of our structural model, the first analysis grounds and enriches our understanding of EPO's effect on patients' outcomes beyond what can be learned from the narrow scope of clinical trials.

We denote the local average causal effect of EPO on observable health outcome c as α_c^{IV}, with our estimation approach captured by

$$y_{ijt}^c = \alpha_0 + \alpha_c^{IV} EPO_{ijt} + X_{ijt} b_c + \varepsilon_{ijt}, \tag{7}$$

where y_{ijt}^c is the health outcome c of patient i treated at facility j in month t and X_{ijt} is a vector of observable patient and facility characteristics, including facility and time fixed effects.[12] The main challenges with identifying the causal effect of EPO stem from reverse causality and simultaneity, which could bias OLS estimates in ambiguous ways. The estimates would be biased upward, for example, if healthier patients receive more EPO, or a downward bias could emerge if unobserved confounds like rapidly declining health lead to both high doses of EPO to combat anemia along with low survival rates.

To overcome these empirical challenges, we use two independent sources of variation in EPO doses within an instrumental variables regression that, under a standard set of assumptions, returns an average treatment effect. First, we use the time-series variation in EPO payments associated with Medicare's payment reform. As Medicare imposed the change uniformly across all providers, rather than targeting specific payment changes to specific facilities, the switch to prospective payments introduced a plausibly exogenous shock to EPO doses due to the change in providers' financial incentives. Second, we use a novel source of variation based on a physiological aspect of anemia management: patients living at higher elevations have higher baseline levels of HGB and therefore require smaller doses of EPO to manage their anemia.

Neither source of variation would represent a valid instrument on its own. Simply using the switch to prospective payments directly would violate the exclusion restriction if trends like updated dialysis

[12]The regression coefficient capturing the differences in the outcome by patient and facility characteristics is denoted b_c to emphasize the fact that this coefficient captures the baseline health differences that do not depend on EPO doses also captured by $b_c(\cdot)$ in the theoretical model.

standards and related medical advances coincide with the payment reform. Similarly, just as elevation directly affects patients' hemoglobin levels, it may also directly affect their health in other ways. We therefore use the interaction of a post-PPS indicator variable and a facility's elevation as an instrument for EPO doses while controlling directly for time fixed effects and elevation in our first- and second-stage regressions. As illustrated by the model, providers considering both their own financial returns as well as their patients' well-being may have responded to the PPS with larger adjustments to EPO doses at lower elevations. Figure 1 supports this prediction, showing that patients at low elevations experienced larger reductions in their EPO doses than patients at higher elevations. As such, we can use the cross-sectional variation from patients' elevations along with the time-series variation induced by the payment reform to identify the causal effect of EPO on outcomes.

Our empirical strategy of interacting one variable with time-series variation and another with cross-sectional variation was first introduced by Bartik (1991) and is closely related to the approaches later used by Card (1995) to study the returns to education, Nunn and Qian (2014) to study the impact of food aid, and Bettinger et al. (2017) to study the effectiveness of online college courses. Adapting this approach to our setting, we have a first-stage specification of

$$EPO_{ijt} = \psi_1 Elevation_j + \psi_2 PPS_t + \psi_3 Elevation_j \times PPS_t + X_{ijt}\Psi_c + \upsilon_{ijt}, \tag{8}$$

where the instrument $Elevation_j \times PPS_t$ varies by facility and time period, allowing us to include month-year fixed effects. By instrumenting for EPO doses with the interaction term, our first stage resembles a difference-in-differences estimation, comparing EPO doses at facilities that typically use less of the drug due to their high elevation with those at lower elevations that typically use more of it, during the fee-for-service era when financial incentives favored higher doses relative to the PPS era when the financial incentives reversed. As outlined in Nunn and Qian (2014), the main distinction between this strategy and a typical difference-in-differences estimation is the continuous treatment variable.

For our second stage to have a causal interpretation, the interaction between a facility's elevation and the timing of Medicare's payment policy must only affect health outcomes through its influence on EPO doses, conditional on the controls. In our setting, this requires that (i) any other mechanism through which elevation affects patients is constant over time and (ii) any other mechanism causing health outcomes to differ before and after the payment reform affects patients uniformly with respect to

their elevation. As discussed above, using elevation as an instrument by itself would mean the reduced-form slope captures both the effect of EPO as well as other plausible mechanisms that affect health outcomes, such as patients living at higher elevations having more-active lifestyles (e.g., hiking and skiing) or elevation having direct consequences for patients' health. By interacting the two variables, however, the reduced-form coefficient measures only how the slope between elevation and outcomes changes when the payment policy changes — the main effect of elevation included in both the first and second stages differences out any other plausible mechanisms that remain constant across the different payment schemes.

Although not directly testable, several pieces of evidence suggest our empirical strategy satisfies these two requirements. In the same spirit as a traditional difference-in-differences estimation, the plot of EPO doses for the first and fifth elevation quintiles in the first panel of Figure 1 shows parallel trends in EPO doses prior to the payment reform.[13] On average, low-elevation patients received higher doses of EPO before prospective payments, with the difference between the two groups remaining constant during this period.[14] After the payment reform, average EPO doses declined in both quintiles, but the decline was much larger for those at lower elevations. The second stage then links the change in EPO to the corresponding health outcomes like transfusions, with the right panel showing a larger increase for patients at lower elevations commensurate with their larger reductions in EPO.

A related threat to identification would be omitted variables that change disproportionately across elevations over time. Although some differences across elevations do exist and change over time, as shown in the balance tables for observable patient characteristics at each elevation quintile in Appendix C, the changes are not systematically moving toward better or worse outcomes. To assess more formally whether changes in unobserved patient characteristics might confound our analysis, we create a composite measure of a patient's health status using the coefficients from an OLS regression of mortality on observable patient characteristics and month-year fixed effects to predict a patient's mortality risk. This predicted mortality variable is likely correlated with unobserved characteristics that affect patients' health, and we can detect changes in the composition of the patient population by testing if predicted mortality changed differentially by elevation after the payment reform. Estimating (8) with predicted

[13] As discussed in Christian and Barrett (2017), non-parallel pre-trends would have suggested our difference-in-differences analog violated the exclusion restriction.

[14] A regression of EPO on facility elevation, a time trend, and the interaction of the two along with patient and facility controls using data prior to prospective payments indicates that the difference in time trends is small and not statistically significant (p=0.5777).

Table 2
Predicted Mortality by Elevation

	(1) Predicted Mortality	(2) Predicted Mortality	(3) Predicted Mortality
Facility Elevation	0.000000182**	0.000000165**	
	(5.95e-08)	(6.05e-08)	
Elevation × PPS	-7.62e-08***	-4.43e-08^{+}	-3.20e-08
	(2.03e-08)	(2.37e-08)	(1.98e-08)
Year-Month FE	0	1	1
Pat/Fac Controls	0	0	0
Facility FE	0	0	1
R-squared	0.000167	0.000431	0.134
Dep. Var. Mean	0.0164	0.0164	0.0164
Observations	10077289	10077289	10077264

Notes: OLS estimates from (8). Dependent variable is predicted mortality. Predicted mortality is the predicted value for each observation using coefficients from a regression of mortality on patient controls and time fixed effects on observations from 2009 and 2010. Time fixed effects are not included in the prediction. PPS is an indicator variable for January 2011 or later. Facility elevation is measured in feet above sea level. An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Standard errors clustered by facility are in parentheses. $^{+}$, *, **, and *** indicate significance at the 10%, 5%, 1%, and 0.1% level, respectively.

mortality as the dependent variable, we show in Table 2 that the differential change by elevation is a precisely estimated zero, suggesting that a changing mix of patients across elevations is unlikely to confound our analysis.[15]

Another violation of the exclusion restriction could come from facilities reinvesting the additional profits they earn from administering less EPO after the payment reform. Facilities at higher elevations received a larger financial gain from Medicare's switch to prospective payments, for instance, and may have reinvested their windfall in ways that improved patient care. As shown in Table 3, however, we do not find evidence of such behavior, as conventional measures of a facility's investment in providing high-quality care, such as the number of patients per staff, the number of patients per station, and the number of infections per patient, did not change across elevations after the payment reform. We further show in Appendix E that we do not find evidence of other potential violations of the exclusion restriction related to the use of drugs other than EPO or changes in dialysis modalities.

[15]As the purpose of these regressions is to assess how predicted mortality correlates with the instrument, and predicted mortality is constructed as a direct function of patient comorbidities and demographics, we omit these patient controls from the specifications in Table 2.

Table 3
Facility Inputs by Elevation

	(1) Nurses Per Technician	(2) Patients Per Employee	(3) Patients Per Station	(4) Employees Per Station	(5) Hosp., Septicemia
Facility Elevation	-0.00000513	-0.0000503	-0.0000399	-0.00000351	-0.000000699***
	(0.0000142)	(0.0000335)	(0.0000588)	(0.0000123)	(0.000000129)
Elevation × PPS	0.00000767	0.0000226	-0.00000837	-0.00000573	0.0000000336
	(0.00000856)	(0.0000231)	(0.0000169)	(0.00000373)	(0.0000000786)
Year-Month FE	1	1	1	1	1
Pat/Fac Controls	1	1	1	1	1
Facility FE	0	0	0	0	0
R-squared	0.226	0.155	0.0866	0.0676	0.00283
Dep. Var. Mean	0.910	5.402	3.988	0.766	0.00939
Observations	242917	254307	256712	256173	10077289

Notes: OLS estimates from (8). Dependent variables in columns (1)–(4) are facility-level ratios. Dependent variable in column (5) is an indicator for hospitalization with a primary diagnosis of septicemia. PPS is an indicator variable for January 2011 or later. Facility elevation is measured in feet above sea level. For columns (1)–(4) an observation is a facility-month. For column (5) an observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include facility elevation, whether the facility is freestanding or hospital-based, and chain ownership status, as well as geographic information about the patient's residence, including state and average income in the ZIP code. For columns (1)–(4) controls are facility-month-level means of the patient-level controls. Standard errors clustered by facility are in parentheses. $^{+}$, *, **, and *** indicate significance at the 10%, 5%, 1%, and 0.1% level, respectively.

With the approach outlined above, our IV estimand recovers the average equilibrium marginal effect of EPO on observable health outcomes, expressed as

$$\alpha_c^{IV} = \frac{\frac{\partial^2 h_c^*}{\partial p \partial a}}{\frac{\partial^2 epo^*}{\partial p \partial a}} = \frac{\alpha_c(x)\frac{\partial^2 epo^*}{\partial p \partial a}}{\frac{\partial^2 epo^*}{\partial p \partial a}} = \alpha_c(x). \tag{9}$$

Using two-stage least squares with a continuous treatment variable, this estimator can be interpreted as a weighted average of the unit causal effects of EPO along the support of EPO doses, with the weights of each incremental level of EPO coming from the size of the complier group at that level. In Appendix E.4, we provide evidence that the compliers with our instrument are drawn from the entire distribution of EPO doses, suggesting our IV estimand is informative about the average causal effect of EPO over the entire patient population.

We can also extend our analysis to directly estimate the heterogeneous effects of EPO by estimating the same IV regression within cells of patients who have identical observable characteristics or, more simply, by interacting observable characteristics with our baseline instrument, allowing us to recover not only the overall local average treatment effect of EPO, but also the heterogeneity in the treatment effect based on differences in patients' observable characteristics like age, gender, and comorbidities (Mogstad and Torgovitsky, 2024). To do so, we estimate the second-stage equation

$$y_{ijt}^c = \alpha_0 + \alpha_{c,1}^{IV} EPO_{ijt} + X_{it} b_{c,X} + EPO_{ijt} \times X_{it} \alpha_{c,X}^{IV} + F_{jt} b_{c,F} + \varepsilon_{ijt}, \tag{10}$$

where F_{jt} represents facility characteristics that may affect the baseline level of health and

$$\mathbb{E}\left[\frac{\partial h_c}{\partial epo^*} | X_{it} = x_{it}\right] = \alpha_{c,1}^{IV} + x_{it}\alpha_{c,X}^{IV}$$

gives the average marginal effect of EPO on patients with observable characteristics X_{it}.[16] Under the

[16] This specification allows the returns from EPO to vary by patient attributes but not by facility characteristics. Different facilities may have production possibilities frontiers that are level-shifts of one another, but the slope does not change. Put differently, if a patient moved from one facility to another, the level of the health outcome y_{ijt}^c could change as facility characteristics enter the $b_c(\cdot)$ function, but the marginal effect of EPO ($\frac{\partial h_c}{\partial epo^*}$) could not. This simplification reflects the physiological and institutional details of anemia treatment: the EPO molecule is the same across providers, and a patient's physiological reaction to a given amount of it will be the same irrespective of which facility administers it. This implicitly assumes the effect of EPO is invariant to elevation. As discussed in Section 2, this assumption is consistent with the medical literature: although elevation shifts baseline health, it is unlikely to influence the direct effects of EPO. Furthermore, in Appendix Section E.3, we estimate a specification instrumenting for EPO within elevation quintiles, finding no differences across them.

assumptions that (i) any heterogeneity in the effect of EPO on observable health outcomes flows only through observable patient characteristics — i.e., there is no unobserved heterogeneity[17] — and (ii) the marginal effect of EPO varies by patient characteristics linearly in an additively separable way, the production function of observable health component c takes the form

$$\alpha_c(x)epo = \left(\alpha_{c,1} + \sum_{x \in X} \alpha_{c,x} x\right) epo, \tag{11}$$

and the estimands $\alpha_{c,1}^{IV}$ and $\alpha_{c,X}^{IV}$ of (10) are the health production function parameters $\alpha_{c,1}$ and $\alpha_{c,x}$, respectively.

To estimate (10), we extend our IV strategy from above using two-stage least squares and treating EPO_{ijt} as an endogenous variable while interacting EPO_{ijt} with all patient attributes in the data. To instrument for these interactions, we use the natural extension of our original instrument by interacting it with each patient attribute and then use these interactions as a new set of instruments. For example, we instrument for the difference in the marginal effect of EPO for men and women using the differential change for men and women after the start of PPS and across elevations. Using analogous instruments for all components of $EPO_{ijt} \times X_{it}$, we estimate (10) and obtain the average marginal effect of EPO for patients with heterogeneous observable characteristics.

5.2. Estimation of Remaining Parameters

After estimating the parameters of the production function for each observable component of health, we then estimate the remaining parameters of the model: the weights providers place on patients' health relative to their own financial returns, β_c, the parameters that translate the observable level of HGB into the health of the patient, $\tau(x)$ and λ, and the standard deviation of the unobservable component of health, σ.

We include HGB levels, cardiac hospitalizations, and mortality for observable components of health, c, with HGB levels serving as a sufficient statistic for anemia management (i.e., as opposed to transfusions, it is the outcome directly targeted by facilities). HGB is likely not a sufficient statistic for overall health effects, however, given the medical literature that suggests that risk factors from EPO arise for

[17]Strictly speaking, we need only assume that providers do not make dosing decisions based on unobserved heterogeneity in the effect of EPO on observable health outcomes. We make the stronger assumption of no unobservable heterogeneity for expositional clarity.

reasons separate from HGB levels. We focus on cardiac hospitalizations, not all-cause hospitalization, given their direct connection to EPO (i.e., a salient side effect of EPO is an elevated risk of heart attacks).

Adding structure to $\tau(x)$, we allow the health-maximizing level of HGB to depend on the patient's baseline HGB, $b_{hgb}(x, a, u)$, and assume it does so in a stepwise fashion. We estimate separate values of τ for patients with baseline HGB less than 10, between 10 and 11, and greater than 11. In other words, we assume the functional form of $\tau(x)$ is

$$\tau(x) = \tau_1 \mathbb{I}[b_{hgb}(x, a, u) < 10] + \tau_2 \mathbb{I}[b_{hgb}(x, a, u) \in [10, 11)] + \tau_3 \mathbb{I}[b_{hgb}(x, a, u) \geq 11].$$

Note that after estimating $\alpha(x)$, $b_{hgb}(x, a, u)$ for a given observation $b_{hgb,ijt}$ is $b_{hgb,ijt} = HGB_{ijt} - \alpha(x_{ijt}) \times EPO_{ijt}$.

We calibrate p, the profit margin on EPO, to match the observed average EPO payment rate and average EPO acquisition costs in facilities' 2010 cost reports separately by chain ownership, with cost estimates by chain reported in Table A14 in Appendix G. Consistent with greater buyer power allowing firms to negotiate substantial discounts from drug manufacturers, we estimate large chains purchase EPO at a much lower cost than independent facilities do.

Finally, we limit the sample to patient-months with a baseline HGB level between 5 and 15 and drop observations with an estimated effect of EPO on HGB α_{HGB} less than 0.005. These restrictions result in dropping 3.4% of the sample.

Using GMM, we separately estimate the model for chain and independent facilities given the differences in their dosing strategies. To match the skewed distribution of EPO doses and heterogeneous responses to the changes in financial incentives, we match the following moments: the overall mean EPO dose; the mean dose for patients with baseline HGB between 10 and 11; the mean dose for patients with baseline HGB greater than 11; the mass between the 25th and 75th percentiles of the empirical EPO distribution; the mass with EPO dose of zero or the maximum dose; and the covariances of EPO doses with financial margin p, baseline HGB, and the marginal effect of EPO on each of HGB, cardiac hospitalizations, and mortality. The covariance moment with financial margin helps identify how providers trade off costs and patient health, while the covariances with the marginal health effects of EPO help

pin down the preference weights providers place on each health dimension.[18]

6. Model Estimation Results

6.1. Health Production Function Parameters

We present results from our first-stage estimates in Table 4, with an F-statistic of 49.1 demonstrating the instrument's relevance. Given the body's physiological response to elevation, EPO doses decrease with elevation in the expected way, but the rate of this decrease falls by over a quarter after the payment reform. Estimates from our preferred specification presented in column (3) indicate that patients at sea level had their average monthly EPO dose reduced by 1400 IUs more than patients living 1000 feet above sea level.

Following the first-stage estimates, we recover the local average treatment effect of EPO on patient outcomes using two-stage least squares. In addition to instrumenting for EPO_{ijt}, we control for several patient covariates, month-year fixed effects, and facility fixed effects, estimating this equation for the main outcomes of interest: HGB levels, blood transfusions, hospitalizations, and mortality.

The results for HGB levels highlight the importance of our empirical strategy. Although randomized controlled trials have shown that EPO increases HGB levels, the OLS results in Table 5 suggest the opposite effect due to the nonrandom assignment of EPO: more-anemic patients with lower HGB levels tend to be prescribed higher doses of EPO, inducing a negative correlation between HGB and EPO if relevant patient attributes are not observed in the data. Our IV strategy corrects for these endogenous treatment decisions, as shown in column (2), where increasing EPO doses by 1000 IUs per month increases a patient's HGB by 0.0210 g/dL, on average, for an implied elasticity of HGB with respect to EPO of 0.09, and confirms the established medical fact that EPO effectively treats anemia. Table 5 also shows results with transfusions as the dependent variable. As with HGB, the OLS coefficient contradicts established medical evidence by suggesting that more EPO is associated with the need for more blood transfusions, whereas correcting for endogenous dosing decisions using our IV strategy reveals that larger EPO doses do indeed reduce the likelihood of receiving a transfusion.

Table 6 shows the primary adverse effects of larger EPO doses — they lead to more hospitalizations

[18]Table A15 reports the empirical means of these moments and shows our ability to match them successfully.

Table 4
First Stage Regression

	(1) EPO	(2) EPO	(3) EPO
Facility Elevation	-0.00477***	-0.00353***	
	(0.000341)	(0.000401)	
Elevation × PPS	0.00144***	0.00133***	0.00140***
	(0.000214)	(0.000203)	(0.000200)
Year-Month FE	1	1	1
Pat/Fac Controls	0	1	1
Facility FE	0	0	1
R-squared	0.0297	0.0835	0.139
Dep. Var. Mean	48.50	48.50	48.50
Observations	10077289	10077289	10077264

Notes: OLS estimates from (8). Dependent variable is monthly EPO dose. EPO doses are censored at the 99th percentile and measured in 1000 IUs. PPS is an indicator variable for January 2011 or later. Facility elevation is measured in feet above sea level. An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include facility elevation (when facility fixed effects are not included), whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. Standard errors clustered by facility are in parentheses. $^{+}$, *, **, and *** indicate significance at the 10%, 5%, 1%, and 0.1% level, respectively.

for cardiac events as well as higher mortality rates. For both all-cause and cardiac hospitalizations, the OLS and IV results suggest a positive correlation with EPO doses, although this effect does not remain statistically significant for all-cause hospitalizations in the IV specification. For mortality, the OLS estimates show a statistically significant, negative correlation with EPO, but the effect becomes positive while remaining statistically significant when we use our instrument. Our IV estimates imply that a 1% increase in EPO raises the rate of cardiac hospitalization by 0.3% and mortality by 0.4%, validating the concerns of policymakers who advocated for the FDA's black box warning and Medicare's

Table 5
The Effect of EPO on Hemoglobin Levels and Transfusions

	HGB		Transfusion	
	(1) OLS	(2) IV	(3) OLS	(4) IV
EPO	-0.00303***	0.0210***	0.000132***	-0.000575***
	(0.0000254)	(0.00545)	(0.00000256)	(0.000154)
Implied Elasticity	-0.0132	0.0916	0.227	-0.989
Year-Month FE	1	1	1	1
Pat/Fac Controls	1	1	1	1
Facility FE	1	1	1	1
Dep. Var. Mean	11.12	11.12	0.0282	0.0282
Observations	8181736	8181736	10077264	10077264
First-Stage F-statistic		33.24		48.97

Notes: OLS and IV estimates from (7). Dependent variable in columns (1)–(2) is hemoglobin. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. Dependent variable in columns (3)–(4) is a binary variable for receiving a blood transfusion. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Implied elasticities are calculated using the reported linear effect of EPO along with the average outcome value and EPO dose reported in Table 1. An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. Standard errors clustered by facility are in parentheses. $^{+}$, *, **, and *** indicate significance at the 10%, 5%, 1%, and 0.1% level, respectively.

adoption of the PPS to reduce the widespread use of EPO among dialysis patients.[19]

As a placebo test, we also estimate (7) with septicemia as the dependent variable. Because septicemia, a severe blood infection resulting from facilities' poor cleaning protocols, has no known relation to EPO, a statistically significant effect of EPO on septicemia would suggest an omitted variable confounds our analysis. As shown in Table 6, however, we do not find a causal effect of EPO on septicemia in our IV specification, providing further reassurance in our identification strategy.

In terms of heterogeneity, Figure 2 shows the distribution of the marginal effects of EPO from

[19]For comparison, Besarab et al. (1998) conducted a randomized controlled trial where dialysis patients with heart conditions were randomized into two EPO treatment arms targeting different hematocrit levels. They found that patients treated most intensely suffered 22% more deaths and 29% more nonfatal heart attacks than those targeting a lower hematocrit level. In comparison to such findings, the substantial effects on mortality and cardiac hospitalizations that we estimate seem reasonable.

Table 6
The Effect of EPO on Hospitalizations and Mortality

	Hosp., Any Cause		Hosp., Cardiac Event		Hosp., Septicemia		Mortality	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
EPO	0.000154***	0.000199	0.0000153***	0.000181+	-0.000000269	0.0000351	-0.000112***	0.000126*
	(0.00000348)	(0.000249)	(0.00000121)	(0.0000943)	(0.000000602)	(0.0000539)	(0.000000893)	(0.0000631)
Implied Elasticity	0.0541	0.0699	0.0274	0.324	-0.00139	0.181	-0.346	0.389
Year-Month FE	1	1	1	1	1	1	1	1
Pat/Fac Controls	1	1	1	1	1	1	1	1
Facility FE	1	1	1	1	1	1	1	1
Dep. Var. Mean	0.138	0.138	0.0271	0.0271	0.00939	0.00939	0.0157	0.0157
Observations	10077264	10077264	10077264	10077264	10077264	10077264	10077264	10077264
First-Stage F-statistic		49.11		49.11		49.11		49.11

Notes: OLS and IV estimates from (7). Dependent variables are binary outcomes. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Implied elasticities are calculated using the reported linear effect of EPO along with the average outcome value and EPO dose reported in Table 1. An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. Standard errors clustered by facility are in parentheses. +, *, **, and *** indicate significance at the 10%, 5%, 1%, and 0.1% level, respectively.

estimating (10) on each of the outcomes we consider for all patient-month observations, and Table 7 reports summary statistics for these distributions.[20] For each outcome, the average of the marginal effects is similar to the local average treatment effects reported in Tables 5 and 6, while the wide variation in patients' responsiveness to EPO has important practical implications. The marginal effect of EPO on HGB, for example, is more than twice as large for a patient one standard deviation more responsive than the mean compared to a patient one standard deviation less responsive. Similarly, the effect of EPO on transfusions is 2.6 times greater, the effect on cardiac hospitalizations is 5.6 times greater, and the effect on mortality is 3.3 times greater. Given the correlation between this treatment effect and the levels of EPO and HGB, the third quartile of the elasticity of HGB with respect to EPO is 6.7 times greater than the first quartile.

Figures A7 through A11 show how specific covariates contribute to the marginal effects of EPO on reducing anemia and unintended side effects. They show, for example, that for patients who are hypertensive, EPO does less to increase their HGB levels or help them avoid transfusions, while doing more to increase their risk of hospitalizations. There is also interesting heterogeneity along demographic dimensions. For example, EPO does less to boost the HGB of males, but also does less to increase their risk of hospitalization or cardiac events.

[20]The coefficients on specific interactions are provided in Appendix F. For each specification, an F-test of the joint significance of the coefficients on the interaction of EPO and patient characteristics has a p-value of less than 0.0001. In the same appendix, we also provide estimates of the CDF of the effects of EPO with confidence intervals indicating the precision of our estimates.

Table 7
Summary Statistics of Estimated Heterogeneous Effects of EPO

	Mean	Std. Dev.	10th perc.	25th perc.	50th perc.	75th perc.	90th perc.
HGB	0.01561**	0.00579**	0.00848+	0.01188**	0.01549**	0.01920**	0.02278**
	(0.00519)	(0.00073)	(0.00442)	(0.00473)	(0.00511)	(0.00557)	(0.00600)
HGB Elasticity	0.08209**	0.10525**	0.00000	0.01590**	0.04728**	0.10687**	0.20379**
	(0.02689)	(0.03042)	(0.00007)	(0.00673)	(0.01687)	(0.03550)	(0.06468)
Transfusions	-0.00056**	0.00025**	-0.00084**	-0.00071**	-0.00057**	-0.00041**	-0.00025
	(0.00018)	(0.00002)	(0.00020)	(0.00018)	(0.00018)	(0.00017)	(0.00017)
Hosp., Any Cause	0.00045	0.00046**	-0.00005	0.00017	0.00043	0.00073**	0.00102**
	(0.00028)	(0.00003)	(0.00027)	(0.00027)	(0.00028)	(0.00029)	(0.00029)
Hosp., Cardiac	0.00023*	0.00016**	0.00004	0.00013	0.00023*	0.00033**	0.00043**
	(0.00011)	(0.00001)	(0.00011)	(0.00011)	(0.00011)	(0.00011)	(0.00012)
Mortality	0.00017**	0.00009**	0.00007	0.00011+	0.00016**	0.00021**	0.00028**
	(0.00001)	(0.00001)	(0.00006)	(0.00007)	(0.00007)	(0.00007)	(0.00008)
Hosp., Septicemia	0.00004	0.00009**	-0.00006	-0.00002	0.00003	0.00008	0.00015*
	(0.00006)	(0.00001)	(0.00006)	(0.00006)	(0.00006)	(0.00006)	(0.00006)

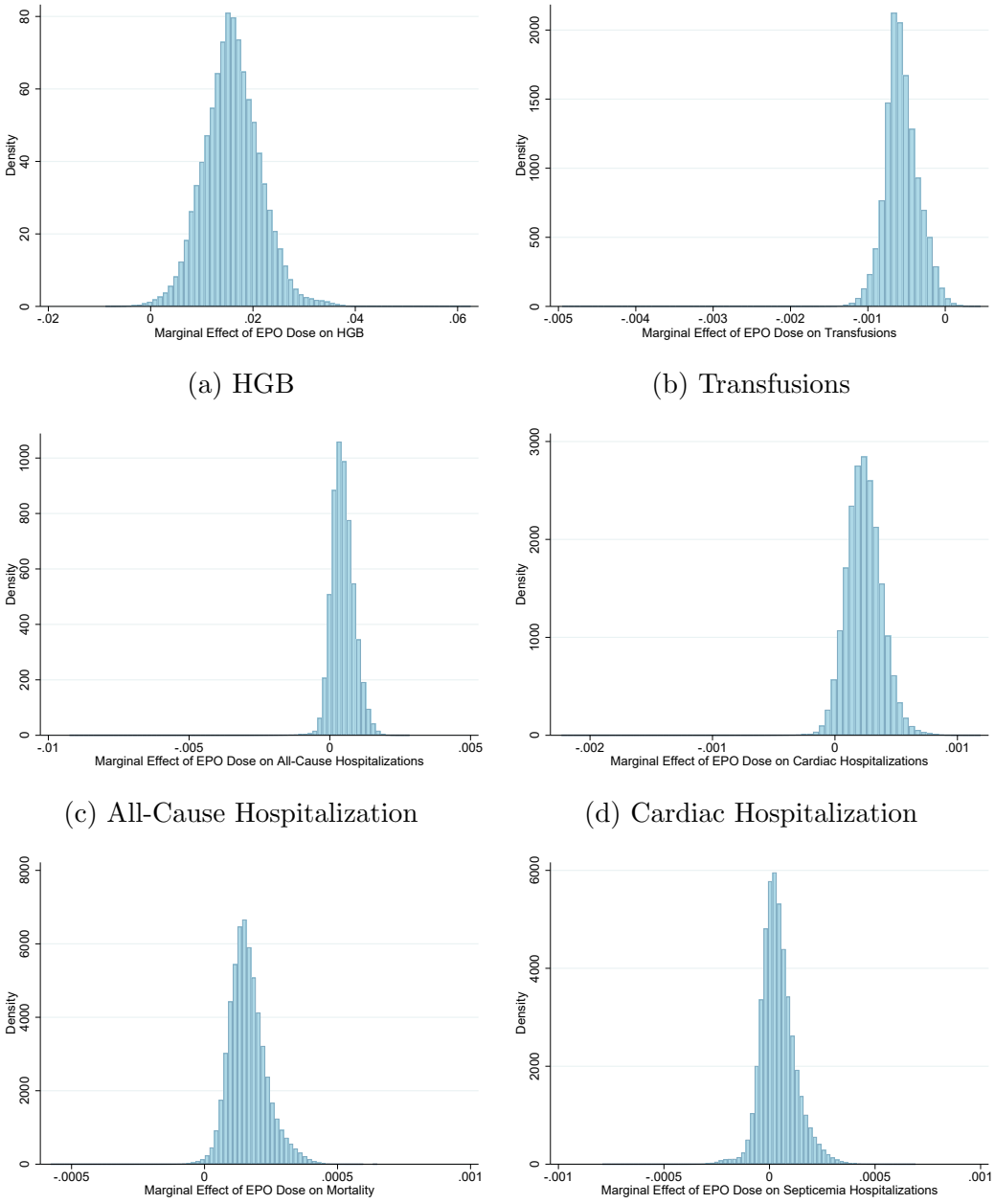
Notes: Summary statistics for α_c using IV estimates of (10). An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. Transfusion and hospitalization variables are binary indicators. HGB elasticity reports the patient-month-level implied elasticity of HGB with respect to EPO using the causal effect reported here along with the observed level of EPO and HGB. Standard errors and p-values are calculated from 200 bootstrapped estimation samples. $^{+}$, *, and ** indicate significance at the 10%, 5%, and 1% level, respectively.

Taken together, our results highlight the clinical tradeoffs associated with using EPO. Although EPO effectively treats patients' anemia, as reflected by higher HGB levels and fewer blood transfusions, these improvements must be weighed against a greater risk of cardiac events and death. In the model, the large degree of heterogeneity in treatment effects illustrates the variation in tradeoffs providers face when weighing the health effects of a dose against the financial considerations: different patients face different risks and benefits from EPO, resulting in a wide range of equilibrium doses and health consequences across patients.

6.2. Comparison to Clinical Trials

Our estimates of EPO's effectiveness add nuance to measures coming from clinical trials. The original Phase III trial, for example, found an average effect per 1000 IUs of EPO of 0.0271 (Eschbach et al., 1989), an effect 30% larger than our estimate from Table 5. The aggregated results from multiple clinical trials in Tonelli et al. (2003) suggest even larger estimates of the average marginal effect, ranging from 0.0338 to 0.0896. Our results show that in practice the average effect may be substantially smaller than what the trials picked up and that focusing only on the average masks important dispersion in

Figure 2
Histogram of Estimated Marginal Effects of EPO



(a) HGB

(b) Transfusions

(c) All-Cause Hospitalization

(d) Cardiac Hospitalization

(e) Mortality

(f) Septicemia Hospitalization

Notes: Estimated marginal effects come from IV estimates of (10). An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. Transfusion and hospitalization variables are binary indicators.

both the risks and benefits of EPO.

One possible source of these discrepancies is the difference between the sample used in the trials and our sample, which can lead to underestimating the negative side-effects of a drug (Abaluck et al., 2025). In the phase III trial, only 23% of patients were hypertensive, 18% were diabetic, and the mean initial HGB was 7 g/dL. By contrast, among those receiving EPO in our sample (i.e., virtually all Medicare patients on dialysis), 91% are hypertensive, 54% are diabetic, and their mean initial HGB is 9.9 g/dL. In other words, the patients included in the original RCTs of EPO are less sick and more anemic than the typical dialysis patient, and our estimates show how these differences contribute to the higher efficacy of EPO measured in clinical trials — hypertension and diabetes each dampen patients' HGB responses to EPO, as shown in Figure A7.

The gap between RCT estimates and real-world effectiveness could also reflect other factors that controlled environments cannot capture adequately. For example, our estimates come from providers' dosing decisions in equilibrium, which incorporate not only financial incentives but also the richer set of tradeoffs weighed when prescribing a particular dose, rather than administrative protocols. Quasi-experimental estimates can complement the evidence from trials by using data that account for these factors, especially when the complier group is more representative of the population than the trial's participants. These distinctions suggest the need for field studies to complement the findings from artificially controlled environments, especially when the drug in question is used as extensively as EPO is and its usage is tied to financial incentives set by Medicare's policy.

6.3. Results for Other Model Parameters

Table 8 presents estimates of the remaining parameters of the model, estimated separately for chain and independent facilities. First, we find plausible HGB targets that are very similar across chain and independent facilities. We find HGB targets close to or within the 10 to 12 g/dL range for all patients, regardless of their baseline HGB level, which was the recommended range at the time of our study. As expected, the HGB target is higher for those with higher baseline HGB, indicating greater kidney function. Also as expected, our estimates of λ are positive, indicating that it is more costly for patients with higher baseline HGB levels to be far away from the targeted range. In terms of magnitude, we estimate that the health cost of having an HGB level 1 g/dL away from one's target is \$64–77 larger

for patients with a baseline HGB level 1 g/dL higher.

The estimates of β_c for each health component give the profit per patient-month facilities are willing to forgo to increase that outcome by one unit.[21] We find that each month chain facilities are willing to pay \$699 to avoid a heart attack, \$21,164 to avoid a death, and \$832 to move from 1 g/dL away from the HGB target to the target, on average.[22] By contrast, we estimate that independent facilities place much greater value on avoiding cardiac hospitalizations at \$10,837, somewhat greater value on avoiding mortality at \$24,913, and similar value on HGB at \$809.[23] This indicates that independent facilities appear more altruistic than chain facilities, with independent facilities generally willing to forgo more profit in order to improve patients' observable health outcomes.

Finally, we estimate a larger value of σ for independent facilities relative to chains, where σ captures the amount of heterogeneity in the effect of EPO on the unobservable health component, expressed in terms of a facility's forgone profits. That is, assuming the variance of the distribution of the effect of EPO on unobservable health is the same across chain and independent facilities, the larger estimate of σ for independent facilities indicates they have more altruism. This is consistent with Eliason et al. (2020), which finds the pre-existing health status of patients at chain and independent facilities to be similar.

6.4. Heterogeneity by Elevation

Although we do not allow the direct effect of EPO on observable health outcomes to vary by elevation, elevation plays a central role in the effect of EPO on overall health. We explore the extent of this heterogeneity in the following ways. First, note that we successfully match the relationship between elevation and EPO despite not directly targeting this moment in estimation, as shown in Figure 3a. In the model, elevation matters for patients' health, particularly the way that baseline HGB varies by elevation. As shown in Figure 3b, the average baseline HGB increases with elevation, meaning that

[21] This can be seen by noting that the derivative of the provider's utility function with respect to total profit, $epo * p$, is 1, while the derivative with respect to health component c is β_c. This implies the provider values both a β_c dollar change in profit and a one unit change in health component c at β_c dollars.

[22] The value of changing a patient's HGB level varies across patients because of their different levels of baseline HGB, which because λ is not equal to zero, means they have different valuations. The coefficient β_{hgb} gives the value of moving from 1 g/dL away from the HGB target to the target for a patient with a baseline HGB of 0 g/dL, while we report the value at the average baseline HGB.

[23] Note that the value chain facilities place on HGB is slightly higher at the average baseline HGB level, which is reported in the text, but lower at a baseline HGB of zero, which is the coefficient β_{hgb} reported in Table 8.

Table 8
Estimates of Weights on Observable Components of Health

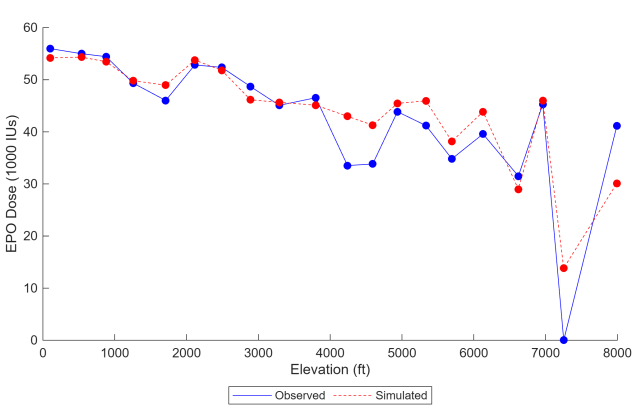
	(1) Chain Facilities	(2) Independent Facilities
τ_1	10.63	10.36
	(0.001)	(0.003)
τ_2	11.38	11.26
	(0.001)	(0.002)
τ_3	12.19	12.04
	(0.001)	(0.002)
λ	77.31	64.50
	(0.592)	(0.399)
β_{HGB}	37.71	146.11
	(4.164)	(4.366)
$\beta_{CardiacHosp.}$	-698.87	-10836.58
	(16.447)	(155.409)
$\beta_{Mortality}$	-21164.10	-24912.56
	(98.234)	(313.894)
σ	2.704	9.251
	(0.147)	(0.062)
Observations	6459685	1445935

Notes: Estimates of model parameters. In column (1), the sample is limited to chain facilities, while it is limited to independent facilities in column (2). An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. The sample is further limited to patient-months with an estimated baseline HGB level between 5 and 15 and an estimated effect of EPO on HGB of greater than 0.005.

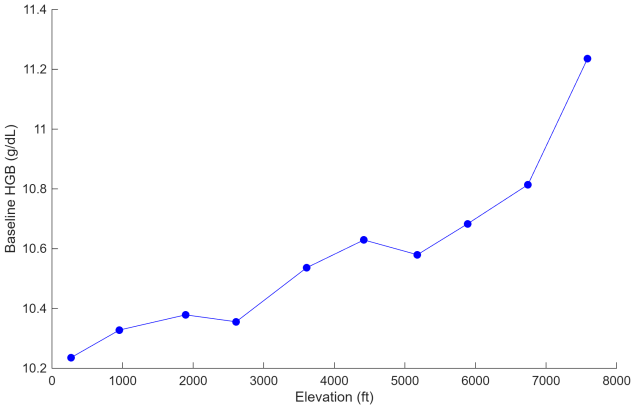
patients at higher elevations on average will have higher HGB targets τ and face a greater health risk from exceeding their target (i.e., λ is positive).

This heterogeneity means that the marginal effect of EPO on total patient health at a given EPO dose decreases with elevation, as shown in Figure 3c. The solid blue line gives the estimated effect of EPO on health at the average EPO dose observed under PPS, with the marginal effect on patients at high elevations significantly more negative on average than for patients at lower elevations. The red line fixes all patients' baseline HGB levels at the average baseline HGB. Eliminating heterogeneity in baseline HGB flattens the relationship between the marginal effect of EPO and elevation, indicating that differences in baseline HGB, rather than differences in other patient characteristics, lead to the estimated heterogeneity.

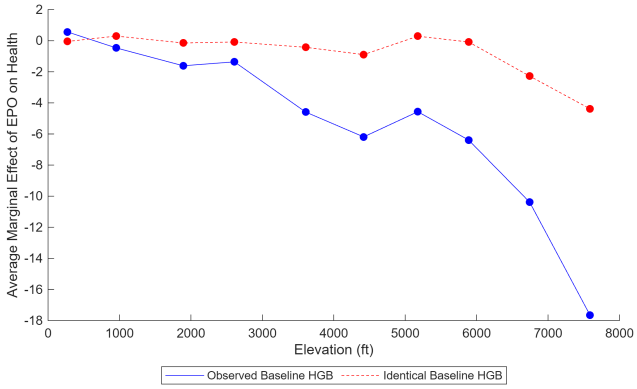
Figure 3
Estimated Characteristics by Elevation



(a) EPO Dose



(b) Baseline HGB



(c) Marginal Effect of EPO on Total Health

Notes: Panel (a) reports the mean observed and model-predicted EPO doses by elevation. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Facility elevation is measured in feet above sea level. Panel (b) reports baseline HGB b_{hgb} by elevation. Baseline HGB is implied by observed HGB and EPO and estimated α_{HGB}. Panel (c) reports the estimated marginal effect of EPO on total health at the average observed EPO dose under PPS (42.47 thousand IUs per month), with the red dashed line fixing all patients' baseline HGB at the average value (10.27 g/dL). Marginal effect is the estimated derivative of the health function H with respect to EPO. For all panels, the sample is broken into 10 bins of equal elevation range, with each point reporting the mean outcome for observations within that bin. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. The sample is further limited to patient-months with an estimated baseline HGB level between 5 and 15 and an estimated effect of EPO on HGB of greater than 0.005.

7. Counterfactual Analysis

Our structural model allows us to assess counterfactual scenarios that illustrate how patient heterogeneity and financial incentives influence the care patients receive and their resulting health outcomes, as well as how they interact with important industry characteristics like chain ownership and drug acquisition costs. Our first counterfactual establishes a point of comparison by considering the health outcomes that could be achieved if facilities simply administered the health-maximizing dose of EPO. We then consider two counterfactuals examining the role of patient heterogeneity in determining EPO doses, one in which all patients receive a uniform dose of EPO set equal to the average health-maximizing amount and another where all patients are treated to achieve the same targeted level of HGB. Finally, we explore the role of dialysis chains in administering EPO by evaluating what would happen if chain facilities had the same level of altruism that we observe for independent facilities and if all facilities acquired EPO at the low cost negotiated by the largest chain. Our main set of counterfactual results are shown in Table 9, with supplemental results in Appendix G.

We first consider the health outcomes that could be achieved if facilities administered a dose to each patient that maximizes their total health, as would be the case if facilities were perfectly altruistic or if Medicare set the per-dose payment equal to its cost. This counterfactual allows us to measure the extent to which financial incentives cause realized health outcomes to diverge from their hypothetical upper bound, with column (2) of Table 9 reporting outcomes under this scenario compared to those observed following the 2011 payment reform reported in column (1). We find that health-maximizing EPO doses are 37% higher, on average, than those observed under PPS, which suggests facilities now undertreat patients.[24] The higher doses in this counterfactual come with adverse tradeoffs — mortality rates increase by 16% and cardiac hospitalization rates increase by 11% — but these negative effects are offset by better controlled anemia and gains in unobserved health, as reflected by higher HGB levels and fewer transfusions. Overall, health-maximizing doses improve total health by an average of $55 per patient per month, where we normalize total health to $0 in the baseline.[25]

To further demonstrate the importance of accounting for patients' heterogeneity when making treat-

[24]The percentage differences are reported in Table A16.

[25]These dollar amounts can be interpreted as the weight independent facilities put on outcomes. If patients value their own health more than the facilities do, these dollar amounts could be scaled up by that difference. For example, if patients value their own health twice as much as an independent facility does, then the total health gains from the health-maximizing dose are twice the values given in Table 9.

Table 9
Counterfactual Outcomes

	(1) Baseline	(2) Health-Maximizing EPO Dose	(3) Uniform EPO Dose	(4) Minimum HGB Target	(5) Chain Altruism Equal to Independent Altruism	(6) Lower Acquisition Cost of EPO
EPO Dose (1000 IUs)	42.47	58.01	58.01	45.23	42.50	43.63
Hemoglobin (g/dL)	10.96	11.17	11.22	11.00	10.95	10.98
Transfusion Rate	0.045	0.036	0.036	0.037	0.039	0.038
All-Cause Hospitalizations	0.132	0.138	0.141	0.139	0.137	0.138
Cardiac Hospitalization	0.024	0.027	0.028	0.028	0.027	0.027
Mortality	0.014	0.016	0.017	0.016	0.016	0.016
Unobserved Health ($)	0	10.691	-74.796	-18.560	87.451	2.202
Total Health ($)	0	55.406	-773.630	-1.372	32.474	8.617

Notes: Outcomes under counterfactual scenarios. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Counterfactual EPO doses are restricted to the observed support of doses. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. Unobserved and total health are measured in dollars and are normalized such that the baseline values are zero. An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. The sample is further limited to patient-months with an estimated baseline HGB level between 5 and 15 and an estimated effect of EPO on HGB of greater than 0.005.

ment decisions, we next consider what would happen if all patients received a uniform EPO dose set to the average health-maximizing level estimated in the previous counterfactual. Column (3) of Table 9 shows how these doses and the corresponding health outcomes compare to those we observe in the data. In this scenario, the average increase in EPO doses over observed doses is the same as the health-maximizing counterfactual by construction, but the variance of EPO doses falls to zero since all patients receive the same amount. As dosing decisions are not tailored to each patient's idiosyncratic needs in this counterfactual, they ignore the diverse tradeoffs incorporated into individualized treatment plans and fail to achieve the gains of the patient-specific health-maximizing dose. Similar to giving individualized health-maximizing doses, patients' anemia is better managed, but acute outcomes such as hospitalizations and mortality become worse. At the same time, the costs in terms of adverse events are larger, while uniform doses impose large unobserved health costs equivalent to nearly $75 per month. The overall impact amounts to a reduction in patient health of almost $800 per month even though doses were set to the average health-maximizing level. Because patients vary in the benefits they receive from EPO, this counterfactual underscores the importance of tailoring doses to each patient's specific needs rather than attempting to bring treatment to a single standard dose for all patients, even if that standard is, on average, a sensible one.

While the previous counterfactual reduced the heterogeneity in the amount of EPO administered to each patient, policymakers may instead want facilities to target EPO doses to reduce heterogeneity in observed health outcomes. For example, the ESRD QIP began penalizing dialysis facilities in 2012 for patients who failed to achieve a minimum HGB level of 10 g/dL (Centers for Medicare & Medicaid Services, 2010). Policies like the QIP that impose uniform targets for all patients ignore the possibility that the ideal outcome may differ across patients depending on their responsiveness to the treatment and sensitivity to its side effects. To evaluate this possibility, we consider the consequences of requiring all patients to achieve a minimum HGB level of 10. Column (4) of Table 9 shows that patients receive 6% more EPO and experience better anemia management in this counterfactual but at the same time accrue health costs that average $1 per month, stemming from both unobserved health costs and an increased risk of adverse events. Although forcing patients to achieve a minimum HGB target is harmful on average, some patients are made better off — the financial incentives facilities face under prospective payments lead them to undertreat patients, whereas requiring a minimum HGB level works to partially offset this incentive. In particular, 61% of patients who previously failed to achieve the HGB target under PPS alone now see health improvements, particularly those who require only small increases in their EPO doses to hit the standard. These patients come closer to their own health-maximizing dose, while others overshoot theirs to reach the minimum level of HGB, worsening their overall health. Allowing heterogeneity in observed health outcomes while also motivating facilities to administer sufficient EPO doses when they have no financial incentive to do so is a key tradeoff policymakers must balance in payment reforms, particularly when financial penalties are tied to health outcomes like in the QIP.

For our final set of counterfactuals, we consider the impact of for-profit chains on the dialysis industry, a long-standing concern for policymakers and patient advocates who contend that these facilities prioritize profits over patients. First, column (5) of Table 9 shows the results of setting chains' level of altruism to that of independent facilities, which reflects one primary reason chains may make different treatment decisions — their fiduciary duty to maximize profits. Because chains now place more weight on total patient health than in our original estimates, we find that average EPO doses increase less than 1% relative to the observed PPS baseline. Despite this small change in doses, the health consequences are substantial, valued at over $32 per patient per month, due to the way in which EPO is reallocated. Here, even though cardiac hospitalizations and mortality rates increase, anemia management and unobserved health improve enough to offset those negative consequences. Compared to the

health-maximizing counterfactual in column (2), setting the altruism of chains to that of independent facilities achieves 59% of patients' maximum possible increase in total health over observed outcomes.

In the last column of Table 9, we reduce the average acquisition cost of EPO for all facilities to that of Fresenius, whose average cost is the lowest we observe in our data. This counterfactual highlights how large chains' purchasing power may influence their treatment decisions, an exercise inspired by recent proposals for Medicare to negotiate directly with manufacturers to procure drugs at a lower cost. Reducing the cost of EPO leads to a 2.7% increase in doses, which results in better overall health for patients but also a slightly higher risk of mortality and cardiac hospitalization. Overall, these higher doses close 16% of the gap between observed patient health and the health-maximizing counterfactual, demonstrating a key tradeoff of allowing chains to provide dialysis: although large chains' lower levels of altruism may harm patients' health, their greater bargaining power over drug manufacturers serves to reduce the conflict between their financial incentives and patients' well-being.

8. Conclusion

We estimate a structural model of dialysis facilities' use of the anti-anemia drug epoetin alfa, historically one of Medicare's largest Part B drug expenses. Using a large payment reform and the elevation of facilities to construct an instrumental variable that identifies our model, we show that doses depend on both the clinical effectiveness of EPO as well as the financial incentives faced by providers, with the substantial heterogeneity in patients' responsiveness to the drug having a large influence on providers' dosing decisions. Although EPO successfully reduces anemia in dialysis patients, facilities must weigh that benefit against an elevated risk of heart attacks and mortality.

We use our model to demonstrate the extent to which the doses administered by providers diverge from those that would maximize a patient's health as well as how allowing providers to tailor treatments to their patients' specific needs results in substantially better outcomes. By contrast, policies that impose uniform health targets or standards of care may harm patients if they prevent providers from taking those specific needs into account. Finally, we show that the large chains that dominate the dialysis industry are less altruistic than independent facilities, but these chains' lower acquisition costs for EPO spur them to use more of the drug, benefiting patients on average.

The large drop in EPO following Medicare's adoption of prospective payments, together with the

wide heterogeneity we estimate for the effect of EPO on patients' health, suggests the need for more work on how providers respond to alternative payment models like the PPS in dialysis and related pay-for-performance initiatives like the QIP. In 2022, alternative payment models accounted for more than 40% of traditional Medicare spending and more than 45% of spending for commercial insurers (Health Care Payment Learning & Action Network., 2023). In light of their growing prevalence, more research focused on the tradeoffs inherent in alternative payment models, both for patients and for other parts of the health care system, would provide valuable guidance for policymakers seeking to restrain costs while at the same time maintaining high standards of care.

References

Abaluck, J., L. Agha, C. Kabrhel, A. Raja, and A. Venkatesh (2016). The Determinants of Productivity in Medical Testing: Intensity and Allocation of Care. *American Economic Review 106*, 3730–3764.

Abaluck, J., L. Agha, and S. Shah (2025). Trials avoid high risk patients and underestimate drug harms. Technical report, National Bureau of Economic Research.

Al-Ubaydli, O., J. A. List, D. LoRe, and D. Suskind (2017a). Scaling for economists: Lessons from the non-adherence problem in the medical literature. *Journal of Economic Perspectives 31*(4), 125–144.

Al-Ubaydli, O., J. A. List, D. LoRe, and D. Suskind (2017b, May). What can we learn from experiments? understanding the threats to the scalability of experimental results. *American Economic Review 107*(5), 282–86.

Allcott, H. (2015). Site selection bias in program evaluation. *The Quarterly Journal of Economics 130*(3), 1117–1165.

Alsan, M., M. Durvasula, H. Gupta, J. Schwartzstein, and H. Williams (2022). Representation and extrapolation: Evidence from clinical trials. *NBER Working Paper*.

Angrist, J. D. and J.-S. Pischke (2009). Mostly harmless econometrics: An empiricist's companion.

Babitt, J. L. and H. Y. Lin (2012). Mechanisms of Anemia in CKD. *Journal of the American Society of Nephrology 23*, 1631–1634.

Badura, K., J. Janc, J. Wasik, S. Gnitecki, S. Skwira, E. Młynarska, J. Rysz, and B. Franczyk (2024). Anemia of chronic kidney disease—a narrative review of its pathophysiology, diagnosis, and management. *Biomedicines 12*(6), 1191.

Baicker, K. and T. Svoronos (2019). Testing the Validity of the Single Interrupted Time Series Design. *NBER Working Paper 26080*.

Baker, S. (2019). The U.S. Health Care System is Full of Monopolies. *Axios*.

Bartik, T. J. (1991). *Who benefits from state and local economic development policies? Kalamazoo, MI: WE Upjohn Institute for Employment Research*.

Bertuzzi, L., P. Eliason, B. Heebsh, R. J. League, R. C. McDevitt, and J. W. Roberts (2023). Gaming and effort in performance pay. Technical report.

Besarab, A., W. K. Bolton, J. K. Browne, J. C. Egrie, A. R. Nissenson, D. M. Okamoto, S. J. Schwab, and D. A. Goodkin (1998). The Effects of Normal as Compared with Low Hematocrit Values in Patients with Cardiac Disease Who Are Receiving Hemodialysis and Epoetin. *New England Journal of Medicine 339*(9), 584–590.

Bettinger, E. P., L. Fox, S. Loeb, and E. S. Taylor (2017). Virtual Classrooms: How Online College Courses Affect Student Success. *American Economic Review 107*(9), 2855–2875.

Bold, T., M. Kimenyi, G. Mwabu, J. Sandefur, et al. (2018). Experimental evidence on scaling up education reforms in kenya. *Journal of Public Economics 168*, 1–20.

Bosman, D., C. Osborne, J. Marsden, I. Macdougall, W. Gardner, and P. Watkins (2002). Erythropoietin response to hypoxia in patients with diabetic autonomic neuropathy and non-diabetic chronic renal failure. *Diabetic Medicine 19*, 65–69.

Bradbury, B. D., M. D. Danese, M. Gleeson, and C. W. Critchlow (2009). Effect of Epoetin Alpha Dose Changes on Hemoglobin and Mortality in Hemodialysis Patients with Hemoglobin Levels Persistently below 11g/dL. *Clinical Journal of the American Society of Nephrology 4*, 630–637.

Brookhart, M., S. Schneeweiss, J. Avorn, B. Bradbury, J. Liu, and W. Winkelmayer (2010, March). Comparative Mortality Risk of Anemia Management Practices in Incident Hemodialysis Patients. *Journal of the American Medical Association 303*(9), 857–864.

Brookhart, M. A., B. D. Bradbury, J. Avorn, S. Schneeweiss, and W. C. Winkelmayer (2011). The Effect of Altitude on Anemia Treatment Responses in Hemodialysis Patients. *American Journal of Epidemiology 173*, 768–777.

Brot-Goldberg, Z. C., A. Chandra, B. R. Handel, and J. T. Kolstad (2017, 04). What does a Deductible Do? The Impact of Cost-Sharing on Health Care Prices, Quantities, and Spending Dynamics. *The Quarterly Journal of Economics 132*(3), 1261–1318.

Card, D. (1995). *Aspects of Labour Market Behavior: Essays in Honor of John Vanderkamp*, Chapter Using Geographic Variation in College Proximity to Estimate the Return to Schooling. Toronto: University of Toronto Press.

Centers for Medicare & Medicaid Services (2010, July). New Bundled Prospective Payment System for End-Stage Renal Disease Facilities Designed to Promote Efficient Care and Improve Patient Outcomes. Technical report, Department of Health & Human Services.

Chan, D. C., M. Gentzkow, and C. Yu (2022). Selection with variation in diagnostic skill: Evidence from radiologists. *The Quarterly Journal of Economics 137*(2), 729–783.

Chandra, A., D. Cutler, and Z. Song (2012). Who Ordered That? The Economics of Treatment Choices in Medical Care. *Handbook of Health Economics 2*, 397–432.

Chandra, A. and D. O. Staiger (2020). Identifying sources of inefficiency in healthcare. *The quarterly journal of economics 135*(2), 785–843.

Chassang, S. and R. Feng (2022). The Cost of Imbalance in Clinical Trials. *NBER Working Paper 30745*.

Chertow, G., J. Liu, K. Monda, D. Gilbertson, M. Brookhart, A. Beaurbrun, W. Winkelmayer, A. Pollock, C. Herzog, A. Ashfaq, T. Sturmer, K. Rothman, B. Bradbury, and A. Collins (2016). Epoetin Alfa and Outcomes in Dialysis amid Regulatory and Payment Reform. *Journal of the American Society of Nephrology 27*, 3129–3138.

Christian, P. and C. Barrett (2017). Revisiting the Effect of Food Aid on Conflict: A Methodological Caution. *World Bank Policy Research Working Paper No. 8171*.

Clemens, J. and J. Gottlieb (2014). Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health? *American Economic Review 104*(4), 1320–1349.

Cuddy, E. and J. Currie (2024). Rules vs. discretion: Treatment of mental illness in us adolescents. *NBER Working Paper No. 27890*.

Currie, J. and W. B. MacLeod (2017). Diagnosing expertise: Human capital, decision making, and performance among physicians. *Journal of labor economics 35*(1), 1–43.

Cutler, D., L. Dafny, and C. Ody (2017). How Does Competition Impact the Quality of Price of Outpatient Service Facilities? A Case Study of the U.S. Dialysis Industry. Working Paper.

Dafny, L. (2005). How Do Hospitals Respond to Price Changes. *American Economic Review 95*, 1525–1547.

Dai, M. (2014). Product Choice Under Price Regulation: Evidence from Out-Patient Dialysis Markets. *International Journal of Industrial Organization 32*, 24–32.

Dai, M. and X. Tang (2015). Regulation and Capacity Competition in Health Care: Evidence from US Dialysis Markets. *The Review of Economics and Statistics 97*(5), 965–982.

DaVita (2005, May). DaVita 2004 Annual Report.

Einav, L., A. Finkelstein, and N. Mahoney (2018). Provider Incentives and Healthcare Costs: Evidence from Long-Term Care Hospitals. *Econometrica 86*(6), 2161–2219.

Eliason, P. (2021). Price Regulation and Market Structure: Evidence from the Dialysis Industry. Working Paper.

Eliason, P. J., P. L. Grieco, R. C. McDevitt, and J. W. Roberts (2018). Strategic Patient Discharge: The Case of Long-Term Care Hospitals. *American Economic Review 108*(11), 3232–65.

Eliason, P. J., B. Heebsh, R. C. McDevitt, and J. W. Roberts (2020). How Acquisitions Affect Firm Behavior and Performance: Evidence from the Dialysis Industry. *The Quarterly Journal of Economics 135*(1), 221–267.

Ellis, R. P. and T. G. McGuire (1986). Provider behavior under prospective reimbursement: Cost sharing and supply. *Journal of health economics 5*(2), 129–151.

Eschbach, J., J. Egrie, M. Downing, J. Browne, and J. Adamson (1987). Correction of the Anemia of End-Stage Renal Disease with Recombinant Human Erythropoietin. *New England Journal of Medicine 316*, 73–78.

Eschbach, J. W., M. H. Abdulhadi, J. K. Browne, B. G. Delano, M. R. Downing, J. C. Egrie, R. W. Evans, E. A. Friedman, S. E. Graber, and N. R. Haley (1989). Recombinant human erythropoietin in anemic patients with end-stage renal disease. results of a phase iii multicenter clinical trial. *Annals of internal medicine 111 12*, 992–1000.

Farag, Y. M. K. et al. (2012). Modulation of platelet activation in chronic kidney disease patients on erythropoiesis-stimulating agents. *Clinical and Applied Thrombosis/Hemostasis 18*(5), 453–461.

Gassmann, M. and M. U. Muckenthaler (2015). Adaptation of iron requirement to hypoxic conditions at high altitude. *Journal of applied physiology 119*(12), 1432–1440.

Gaynor, M., N. Mehta, and S. Richards-Shubik (2023). Optimal Contracting with Altruistic Agents: Medicare Payments for Dialysis Drugs. *American Economic Review 113*(6), 1530–1571.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020). Bartik instruments: What, when, why, and how. *American Economic Review 110*(8), 2586–2624.

Green, A., N. Trivedi, J. Hsu, N. Yu, P. Bach, and S. Chimonas (2022). Despite the fda's five-year plan, black patients remain inadequately represented in clinical trials for drugs. *Health Affairs 41*, 368–374.

Grieco, P. and R. C. McDevitt (2017). Productivity and Quality in Health Care: Evidence from the Dialysis Industry. *The Review of Economic Studies 84*(3), 1071–1105.

Health Care Payment Learning & Action Network. (2023). Apm measurement: Progress of alternative payment models.

Hirth, R., M. Turenne, A. Wilk, J. Wheeler, K. Sleeman, W. Zhang, M. Paul, T. Nahra, and J. Messana (2014). Blood Transfusion Practices in Dialysis Patients in a Dynamic Regulatory Environment. *American Journal of Kidney Disease 64*, 616–621.

Kalantar-Zadeh, K. (2017). History of Erythropoiesis-Stimulating Agents, the Development of Biosimilars, and the Future of Anemia Treatment in Nephrology. *American Journal of Nephrology 45*, 235–247.

Kalantar-Zadeh, K., G. Lee, J. Miller, J. Jing, J. Robertson, and C. Kovesdy (2009). Predictors of Hyporesponsiveness to Erythropoiesis-Stimulating Agents in Hemodialysis Patients. *American Journal of Kidney Disease 53*, 823–834.

Kidney Disease: Improving Global Outcomes (KDIGO) Anemia Work Group (2012). Kdigo clinical practice guideline for anemia in chronic kidney disease. *Kidney International Supplements 2*, 279–335.

Krapf, R. and H. N. Hulter (2009, February). Arterial hypertension induced by erythropoietin and erythropoiesis-stimulating agents (esa). *Clinical Journal of the American Society of Nephrology 4*(2), 470–480.

League, R. J., P. Eliason, R. C. McDevitt, J. W. Roberts, and H. Wong (2022). Variability in prices paid for hemodialysis by employer-sponsored insurance in the us from 2012 to 2019. *JAMA Network Open 5*(2), e220562–e220562.

Lin, E. (2021, 03). The Cost of Transferring Dialysis Care From the Employer-Based Market to Medicare. *JAMA Network Open 4*(3), e212113–e212113.

McGuire, T. (2000). "physician agency" chapter 9 in handbook of health economics, vol i.

Mogstad, M. and A. Torgovitsky (2024). Instrumental variables with unobserved heterogeneity in treatment effects. In *Handbook of labor economics*, Volume 5, pp. 1–114. Elsevier.

Mueller, T. (2023). *How to Make a Killing: Blood, Death and Dollars in American Medicine*. W. W. Norton & Company.

Murthy, V. H., H. M. Krumholz, and C. P. Gross (2004). Participation in cancer clinical trials: race-, sex-, and age-based disparities. *Jama 291*(22), 2720–2726.

Nakhoul, G. and J. F. Simon (2016). Anemia of chronic kidney disease: Treat it, but not too aggressively. *Cleve Clin J Med 83*(8), 613–624.

Nunn, N. and N. Qian (2014). US Food Aid and Civil Conflict. *American Economic Review 104*(6), 1630–1666.

Oostrom, T. (2022). Funding of Clinical Trials and Reported Drug Efficacy. *Working Paper*.

Sibbel, S., B. J. Maroni, and S. M. Brunelli (2017). The Effect of Altitude on Erythropoiesis-Stimulating Agent Dose, Hemoglobin Level, and Mortality in Hemodialysis Patients. *Journal of Nephrology 30*, 821–829.

Silver, D. (2021). Haste or waste? peer pressure and productivity in the emergency department. *The Review of Economic Studies 88*(3), 1385–1417.

Singh, A., L. Szczech, K. Tang, H. Barnhart, S. Sapp, M. Wolfson, and D. Reddan (2006, November). Correction of Anemia with Epoeitin Alfa in Chronic Kidney Disease. *The New England Journal of Medicine 355*, 2085–2098.

Thamer, M., Y. Zhang, J. Kaufman, D. Cotter, F. Dong, and M. A. Hernan (2007). Dialysis Facility Ownership and Epoetin Dosing in Patients Receiving Hemodialysis. *Jama 297*(15), 1667–1674.

Thamer, M., Y. Zhang, D. Lai, O. Kshirsagar, and D. Cotter (2013). Influence of Safety Warnings on ESA Prescribing among Dialysis Patients Using an Interrupted Time Series. *BMC Nephrology 14*.

Tonelli, M., W. Winkelmayer, K. Jindal, W. Owen, and B. Manns (2003). The cost-effectiveness of maintaining higher hemoglobin targets with erythropoietin in hemodialysis patients. *Kidney International 64*, 295–304.

United States Renal Data System (2019). USRDS 2019 Annual Data Report: Epidemiology of Kidney Disease in the United States. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases. Bethesda, MD.

U.S. Geological Survey Center for Earth Resources Observation and Science (2014). GTOPO30 Global 30 Arc-Second Elevation.

U.S. Government Accountability Office (2012, October). Medicare Part B Drug Spending. Publication No. GAO-13-46R, Retrieved from `https://www.gao.gov/assets/650/649459.pdf`.

Valderrabano, F. (2000). Quality of Life Benefits of Early Anemia Treatment. *Nephrology Dialysis Transplant 15*, 23–28.

van der Putten, K., B. Braam, K. E. Jie, and C. A. J. M. Gaillard (2008). Mechanisms of disease: erythropoietin resistance in patients with both heart and kidney failure. *Nature Clinical Practice Nephrology 4*(1), 47–57.

Wang, V., C. Coffman, L. Sanders, S. Lee, R. Hirth, and M. Maciejewski (2018). Medicare's New Prospective Payment System on Facility Provision of Peritoneal Dialysis. *Clinical Journal of the American Society of Nephrology 13*, 1833–1841.

Whoriskey, P. (2012, July). Anemia Drug Made Billions, but at What Cost. *The Washington Post*.

Wilson, N. E. (2016a). For-Profit Status and Industry Evolution in Health Care Markets: Evidence from the Dialysis Industry. *International Journal of Health Economics and Management 16*(4), 297–319.

Wilson, N. E. (2016b). Market Structure as a Determinant of Patient Care Quality. *American Journal of Health Economics 2*(2), 241–271.

Winkelmayer, W. C., J. Liu, and M. Brookhart (2009). Altitude and All-Cause Mortality in Incident Dialysis Patients. *JAMA 301*(5), 508–512. 10.1001/jama.2009.84.

Xia, X., W. Deng, P. J. Eliason, R. J. League, R. C. McDevitt, J. W. Roberts, and H. Wong (2025). Financial ties, market structure, commercial prices, and medical director compensation in dialysis. In *JAMA Health Forum*, Volume 6, pp. e252659–e252659. American Medical Association.

Zhang, Q., M. Thamer, O. Kshirsagar, and Y. Zhang (2017). Impact of the end stage renal disease prospective payment system on the use of peritoneal dialysis. *Kidney International Reports 2*(3), 350–358.

Zhang, Y., M. Thamer, K. Stefanik, J. Maukman, and D. Cotter (2004). Epoetin Requirements Predict Mortality in Hemodialysis Patients. *American Journal of Kidney Disease 45*, 866–876.

Appendix: For Online Publication

The following appendices provide additional robustness checks, analyses, and details on our data.

Appendix A illustrates the robustness of our results to differences in the timing of PPS adoption.

Appendix B shows that neither the black box warnings nor the QIP can explain the patterns we observe for EPO doses.

Appendix C contains additional summary statistics by quintile of facility elevation.

Appendix D shows that our results are robust to a possible anticipatory response by providers.

Appendix E provides more results supporting our identifying assumptions.

Appendix F reports additional parameter estimates.

Appendix G presents additional results from the estimation of our structural model.

A. Differences in Timing of PPS Adoption

The PPS program allowed providers to gradually transition to the bundling of injectable drugs with the dialysis session such that this bundled payment comprised 25% of payments in 2011, 50% in 2012, 75% in 2013, and 100% in 2014. Alternatively, facilities could exercise a one-time option to opt in by November 2010 and immediately receive all payments under the PPS in 2011. Here, we present results showing the vast majority of providers chose to immediately transition to the new PPS and our baseline results are very similar to the results if we use only the subset of immediate-adopters.

First, we attempt to determine within our data the number of facilities that chose to immediately transition to the PPS by documenting whether a facility receives any positive payments for an injectable drug administered to a patient, which we view as a conservative measure of whether a facility has not fully adopted the PPS. We find that whereas more than 99.9% of facilities received payments for an injectable drug in each year prior to 2011, only 7.7% of facilities did afterwards, implying that over 92% of facilities immediately transitioned to the PPS based on this measure. The number increases to the point of full adoption by 2014, with independently owned facilities comprising 83.4% of those that transitioned gradually.

Next, we compare EPO use and patient outcomes by facility according to whether the facility immediately transitioned to the PPS ("Immediate") or not ("Gradual"). Table A1 shows this comparison using data from 2010. We find that patient outcomes are quite similar across these facilities, while those that opted for a gradual transition tended to use less EPO, primarily because most of the facilities that transitioned gradually were independent, which use less EPO on average. Furthermore, we do not find large elevation differences between the facilities. These facts, along with the small number of facilities that did not immediately transition, provide reassurance that selection bias does not undermine our estimates.

Nonetheless, we re-estimate our baseline results for the average treatment effects of EPO using only the sample of facilities that immediately transition to the PPS. The results, shown in Table A2, demonstrate that our baseline results are robust to focusing solely on this set of facilities.

Table A1
Summary Statistics by Immediate Transition to PPS

	PPS Without Transition?		
	Opts Out	Opts In	Total
Facility Characteristics			
Facility Elevation (ft)	644.9	639.6	641.3
Independent Ownership	0.835	0.152	0.209
EPO Use			
EPO Dose (1000 IUs)	39.13	59.09	57.02
Receives Any EPO	0.550	0.796	0.769
Health Outcomes			
Hemoglobin (g/dL)	11.26	11.33	11.32
Transfusions	0.032	0.026	0.026
Mortality	0.017	0.016	0.016
Hospitalizations			
Any Cause	0.1484	0.1412	0.1407
Cardiac Event	0.0282	0.0282	0.0280
Septicemia	0.0113	0.0092	0.0092
Patient-Months	167,827	2,282,122	2,485,214

Notes: An observation is a patient-month. Sample consists of observations from January to December 2010 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later and who are treated at a facility that does not permanently close before 2011. Facilities that opt out of PPS without transition are those for which positive payments for injectable drugs are observed in 2011 or 2012. Facilities that opt into PPS without transition are those for which no payments for injectable drugs are observed in 2011 or 2012 but which received other payments. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. Facility elevation is measured in feet above sea level. Transfusion and hospitalization variables are binary indicators.

Table A2
Baseline Results Using only Facilities that Immediately Transition to PPS

	HGB	Transfusion	Mortality	Hosp., Any Cause	Hosp., Cardiac Event	Hosp., Septicemia
EPO	0.0284***	-0.000616**	0.000151*	0.000227	0.000211+	0.0000283
	(0.00820)	(0.000189)	(0.0000756)	(0.000305)	(0.000117)	(0.0000666)
Year-Month FE	1	1	1	1	1	1
Pat/Fac Controls	1	1	1	1	1	1
Facility FE	1	1	1	1	1	1
Dep. Var. Mean	11.13	0.0279	0.0157	0.138	0.0273	0.00932
Observations	7609185	9249810	9249810	9249810	9249810	9249810
First-Stage F-statistic	20.78	34.27	34.27	34.27	34.27	34.27

Notes: IV estimates from (7). Dependent variable in column (1) is hemoglobin. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. Dependent variables in columns (2)–(6) are binary outcomes. EPO doses are censored at the 99th percentile and measured in 1000 IUs. An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later and who are treated at a facility that neither permanently closes before 2011 nor is observed to receive separate payment for injectable drugs in 2011 or later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. Standard errors clustered by facility are in parentheses. +, *, **, and *** indicate significance at the 10%, 5%, 1%, and 0.1% level, respectively.

B. The Effect of Black Box Warnings & QIP on EPO

Although the FDA's updated black box warning for EPO and Medicare's introduction of the QIP for dialysis facilities occurred around the same time as the payment reform, we present evidence that neither contributed meaningfully to the decline in EPO doses shown in the paper. For the black box warning, four institutional details suggest that it did not cause the decrease in EPO around 2011. First, we show in Appendix E that other injectable drugs, which did not receive black box warnings, follow a pattern similar to EPO's after the PPS. Second, as we discuss in Section 2.1, the FDA has issued two black box warnings for EPO, both of which recommend providers use EPO more judiciously, but the evolution of EPO doses in Figure A1 shows that they did not change following the first black box warning in 2007, an instance when the label changed but financial incentives did not. Third, the decline in EPO begins in October 2010, eight months before the black box warning update, and it is unclear why providers would have changed their behavior in anticipation of the new black box warning even if they had been aware of the FDA's looming decision given that they did not change their behavior following the first black box warning. Finally, as shown by Figure A2 a coincidental drop in EPO use stemmed from one large chain that renegotiated its contract with drug supplier Amgen in mid-2011, as other chains and independent facilities do not exhibit the same patterns for EPO doses.

The large dialysis chains DaVita and Fresenius have at times partnered with Amgen, a leading producer of erythropoietin stimulating agents (ESAs), to make administering drugs such as EPO more profitable. In 2011, DaVita entered into a sourcing and supply agreement with Amgen, providing DaVita with discounts and rebates for Amgen's two ESAs, EPOGEN and Aranesp (DaVita Amgen Agreement 2011). In return, DaVita agreed to purchase at least 90% of its ESAs from Amgen. This 2011 contract ran through 2018 and was renewed in 2017 to extend through 2022 (DaVita Amgen Agreement 2017). Fresenius entered into a similar sourcing and supply agreement with Amgen in 2006, extending to 2011 (Fresenius Amgen Agreement 2006). Fresenius' contract lacked minimum purchase commitments, but did secure discounts for EPOGEN and Aranesp. Our understanding is that Fresenius now has year-to-year contracts with Amgen.

The distinct drop in average HGB levels in mid-2011 corresponds to the renegotiation of multiple large chains' contracts with Amgen, the monopoly supplier of EPO at the time. We see that the sharp drop in EPO and HGB levels in mid-2011 occurred only for patients at one of these large chains. This

provides further evidence that the cause of the discrete drop in EPO and HGB after the initial response to the payment reform was likely not the FDA black box warning but rather the renegotiation of this chain's supply agreement with Amgen. Because the contract renegotiation occurred at the same time as the PPS was implemented, the renegotiation likely reflected a change in this particular chain's strategy following the PPS. If this is the case, then the drop in EPO and HGB occurring in mid-2011 would be attributable to the PPS, with the delay highlighting the sticky nature of chains' supply agreements.

The other policy change around the start of the PPS was the QIP. As we explain in Section 2.2, Medicare instituted the QIP along with the PPS to provide facilities with incentives for maintaining high-quality care while still restraining health care costs. In contrast to the PPS, which focuses on cost containment, the QIP aims to promote a high standard of care by reducing payments to poorly performing facilities.

To implement the QIP, each year Medicare announces the various performance measures that will comprise a facility's Total Performance Score (TPS). Facilities whose scores fall short of the benchmark that year face a reduction of their Medicare payments of between 0.5%–2.0%, depending on the extent of the shortfall. During the sample period for our paper, Medicare used three clinical measures to construct the TPS: the percentage of patients with (i) HGB below 10 g/dL, (ii) HGB above 12g/dL, and (iii) urea reduction ratio (URR) above 0.65. For the first year of the QIP in 2012, Medicare used the facility's performance on these measures in 2010 to construct the TPS. For 2013 and 2014, only the latter two measures were used (based on facility performance in 2011 and 2012, respectively), with Medicare dropping low HGB levels as a criteria. The QIP also included a measure of vascular access in the TPS for 2014, although vascular access has no relation to EPO or other injectable drugs included in the payment reform, so we do not discuss it here.

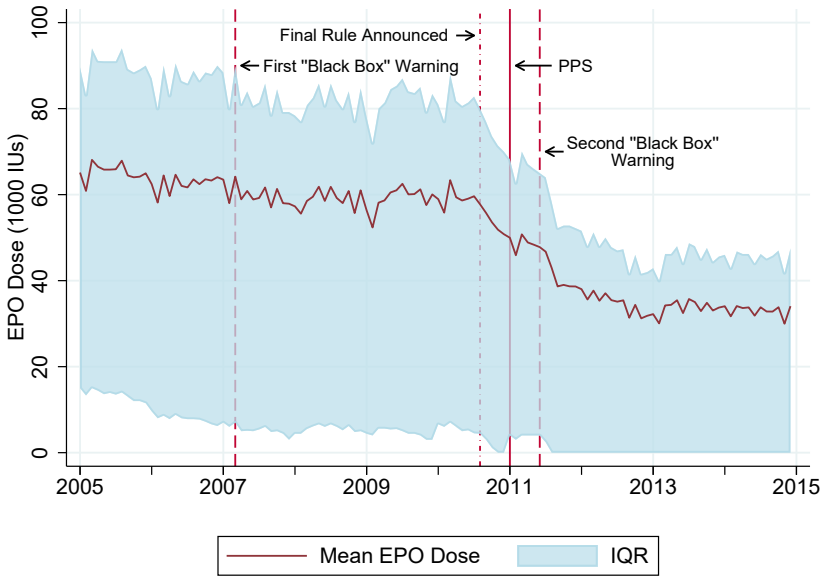
Although Medicare introduced the QIP to discipline facilities' behavior, Figure A3 shows that it did not cause the decline in EPO doses during this period—if anything, the QIP likely makes our estimate of the PPS's impact on EPO doses a conservative one. In Figure A3a, which shows the percentage of patients with HGB greater than 12 g/dL, we see no change in trend following the announcement of this performance measure in 2010. Because EPO directly affects patients' HGB levels, the fact that the trend in the proportion of patients with high HGB levels remained constant after facilities began receiving penalties suggests this standard had little impact on dosing decisions.

Figure A3b shows the percentage of patients with HGB below 10 g/dL.[26] Again, facilities did not respond to the metric's introduction, with the trend remaining constant throughout 2010, although we do see evidence consistent with facilities responding to the metric's removal in 2011. The sharp rise in patients with HGB less than 10 g/dL after Medicare removed this metric from the QIP suggests that (i) our estimates of the PPS's impact on EPO and outcomes are potentially understated, because facilities may have continued giving EPO to low-HGB patients to avoid QIP penalties, and (ii) direct financial incentives from EPO payments predominately dictate facilities' dosing decisions, as facilities cut EPO doses to reduce their drug costs immediately upon Medicare's removal of the low-HGB guardrails.

In short, although the black box warning in 2011 and the QIP performance measures applied to 2010–2012 could have potentially confounded our analysis of the payment reform's effect on EPO doses, we find little evidence that they did, and, if anything, they suggest our results may be conservative. Moreover, because Medicare introduced the QIP in conjunction with the PPS, any potential confounding from the QIP would simply add nuance to our interpretation of the reforms rather than undermine our main findings. That is, we find that the financial incentives from the payment reform had a much stronger influence on facility behavior than the penalties from the QIP did, which provides valuable insights to policymakers aiming to restrain health care costs while maintaining high standards for care. We consider the full effects of the QIP in Bertuzzi et al. (2023).

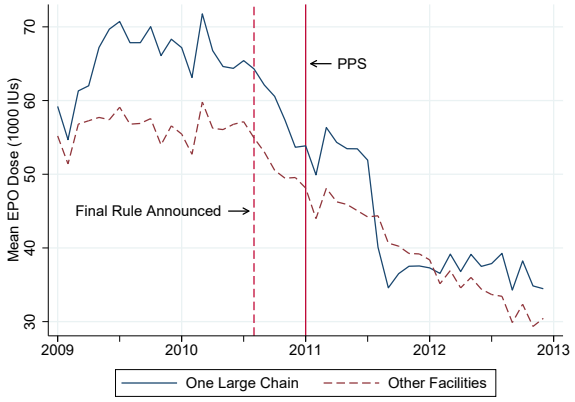
[26]The removal of the measure relating to the percentage of patients with HGB below 10 g/dL was announced in July 2011 and retroactively applied to the performance year beginning January 2011. This means that the TPS calculated using facilities' performances from January to December of 2011 did not include the percentage of patients with HGB below 10 g/dL, but facilities did not learn that this measure would not be used until midway through the year. This proposed rule change was finalized by Medicare in November 2011.

Figure A1
Monthly EPO Doses Over Time with Black Box Warnings

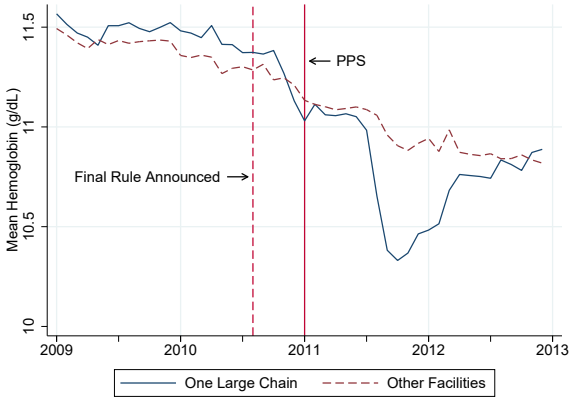


Notes: Figure reports the mean monthly EPO dose and shades the interquartile range of monthly EPO dose. An observation is a patient-month. Sample consists of observations from January 2005 to December 2014 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Vertical long-dashed lines indicate the release of official warnings from the FDA about the safety of high EPO doses. The solid vertical line indicates the start of the PPS in January 2011, while the dot-dashed vertical line indicates the announcement of the final rule for the PPS.

Figure A2
EPO Doses and HGB by Facility Ownership



(a) Monthly EPO Dose for One Large Chain and Other Facilities' Patients



(b) Mean HGB for One Large Chain and Other Facilities' Patients

Notes: An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. The solid vertical line indicates the start of the PPS in January 2011, while the dashed vertical line indicates the announcement of the final rule for the PPS.

Figure A3
QIP HGB Performance Measures

(a) HGB > 12 g/dL Over Time

(b) HGB < 10 g/dL Over Time

Notes: Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. An observation is a patient-month. Sample consists of observations from January 2005 to December 2014 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Vertical lines indicate the introduction and removal of the QIP performance measure.

C. Summary Statistics by Elevation

We provide additional summary statistics for our data by quintile of facility elevation in three tables. The first pools patients across years from 2009 to 2012. The next two show summary statistics for 2009 and 2012 separately. Patients at higher elevations tend to be somewhat less healthy than those at lower elevations, but these differences do not change following the start of the PPS. We do, however, see EPO doses and outcomes change differentially by elevation, providing descriptive evidence that the policy had different effects depending on a patient's elevation.

Table A3
Patient Descriptive Statistics by Elevation

	Elevation Quintile					
	First	Second	Third	Fourth	Fifth	Total
Patient Characteristics						
Predicted Mortality	0.016	0.015	0.016	0.017	0.017	0.016
Age (Years)	63.41	63.60	62.91	63.53	63.57	63.40
Months with ESRD	45.59	45.35	45.72	45.49	43.22	45.08
Black	0.447	0.440	0.452	0.375	0.211	0.385
Male	0.553	0.548	0.545	0.551	0.562	0.552
Diabetic	0.526	0.534	0.536	0.544	0.560	0.540
Hypertensive	0.910	0.906	0.909	0.905	0.900	0.906
Incident Hemoglobin	9.755	9.786	9.806	9.901	10.018	9.853
Facility Characteristics						
Facility Elevation (ft)	29.4	143.7	436.1	713.5	1875.9	638.1
Independent Ownership	0.185	0.183	0.177	0.231	0.208	0.197
EPO Use						
EPO Dose (1000 IUs)	51.50	50.24	50.94	46.84	42.90	48.50
Receives Any EPO	0.791	0.784	0.779	0.725	0.694	0.755
Health Outcomes						
Hemoglobin (g/dL)	11.11	11.11	11.12	11.12	11.16	11.12
Transfusions	0.030	0.028	0.028	0.028	0.027	0.028
Mortality	0.015	0.015	0.015	0.016	0.017	0.016
Hospitalizations						
Any Cause	0.1406	0.1382	0.1355	0.1418	0.1340	0.1380
Cardiac Event	0.0280	0.0281	0.0268	0.0280	0.0248	0.0271
Septicemia	0.0097	0.0095	0.0091	0.0095	0.0090	0.0094
Unique Patients	102,897	99,507	102,182	103,307	103,770	461,477
Patient-Months	2,043,637	1,989,978	2,033,229	2,000,408	2,010,037	10,077,289

Notes: An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Predicted mortality is the predicted value for each observation using coefficients from a regression of mortality on patient controls and time fixed effects on observations from 2009 and 2010. Time fixed effects are not included in the prediction. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. Facility elevation is measured in feet above sea level. The cut points between elevation quintiles are 73, 260, 599, and 870 feet above sea level.

Table A4
Patient Descriptive Statistics by Elevation, 2009

	Elevation Quintile					
	First	Second	Third	Fourth	Fifth	Total
Patient Characteristics						
Predicted Mortality	0.016	0.015	0.016	0.017	0.017	0.016
Age (Years)	63.44	63.57	62.98	63.65	63.83	63.49
Months with ESRD	42.29	42.25	42.39	42.53	40.03	41.90
Black	0.446	0.438	0.447	0.370	0.207	0.382
Male	0.550	0.546	0.543	0.549	0.559	0.549
Diabetic	0.510	0.524	0.524	0.531	0.549	0.528
Hypertensive	0.908	0.905	0.910	0.904	0.899	0.905
Incident Hemoglobin	9.836	9.855	9.866	9.975	10.094	9.925
Facility Characteristics						
Facility Elevation (ft)	29.8	143.3	437.8	714.2	1868.8	638.0
Independent Ownership	0.199	0.202	0.195	0.267	0.229	0.218
EPO Use						
EPO Dose (1000 IUs)	63.28	61.73	62.19	55.73	52.35	59.07
Receives Any EPO	0.813	0.802	0.795	0.732	0.713	0.771
Health Outcomes						
Hemoglobin (g/dL)	11.46	11.45	11.44	11.45	11.46	11.45
Transfusions	0.026	0.025	0.025	0.026	0.024	0.025
Mortality	0.016	0.016	0.017	0.018	0.017	0.017
Hospitalizations						
Any Cause	0.1471	0.1446	0.1420	0.1463	0.1391	0.1438
Cardiac Event	0.0307	0.0303	0.0289	0.0300	0.0267	0.0293
Septicemia	0.0093	0.0091	0.0088	0.0089	0.0084	0.0089
Unique Patients	54,576	52,150	54,661	53,701	54,001	256,504
Patient-Months	477,695	457,844	478,139	467,866	468,898	2,350,442

Notes: An observation is a patient-month. Sample consists of observations from January to December 2009 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Predicted mortality is the predicted value for each observation using coefficients from a regression of mortality on patient controls and time fixed effects on observations from 2009 and 2010. Time fixed effects are not included in the prediction. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. Transfusion and hospitalization variables are binary indicators. Facility elevation is measured in feet above sea level. The cut points between elevation quintiles are 73, 260, 599, and 870 feet above sea level.

Table A5
Patient Descriptive Statistics by Elevation, 2012

	Elevation Quintile					
	First	Second	Third	Fourth	Fifth	Total
Patient Characteristics						
Predicted Mortality	0.016	0.016	0.016	0.017	0.017	0.016
Age (Years)	63.37	63.63	62.85	63.35	63.33	63.31
Months with ESRD	48.98	48.68	49.02	48.59	46.44	48.34
Black	0.448	0.443	0.454	0.379	0.213	0.388
Male	0.556	0.551	0.546	0.554	0.565	0.554
Diabetic	0.538	0.542	0.546	0.555	0.569	0.550
Hypertensive	0.911	0.908	0.909	0.906	0.902	0.907
Incident Hemoglobin	9.664	9.710	9.737	9.819	9.935	9.772
Facility Characteristics						
Facility Elevation (ft)	29.2	144.3	434.4	713.6	1886.7	637.2
Independent Ownership	0.172	0.161	0.150	0.197	0.184	0.173
EPO Use						
EPO Dose (1000 IUs)	36.71	36.11	36.75	34.27	30.43	34.87
Receives Any EPO	0.759	0.761	0.751	0.708	0.662	0.728
Health Outcomes						
Hemoglobin (g/dL)	10.79	10.81	10.82	10.83	10.89	10.83
Transfusions	0.033	0.030	0.0306	0.030	0.029	0.030
Mortality	0.015	0.014	0.015	0.015	0.015	0.015
Hospitalizations						
Any Cause	0.1344	0.1305	0.1283	0.1348	0.1275	0.1311
Cardiac Event	0.0257	0.0258	0.0246	0.0256	0.0227	0.0249
Septicemia	0.0103	0.0100	0.0094	0.0099	0.0094	0.0098
Unique Patients	60,055	58,219	58,652	58,026	58,970	280,751
Patient-Months	543,541	528,788	531,440	518,537	527,525	2,649,831

Notes: An observation is a patient-month. Sample consists of observations from January to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Predicted mortality is the predicted value for each observation using coefficients from a regression of mortality on patient controls and time fixed effects on observations from 2009 and 2010. Time fixed effects are not included in the prediction. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. Transfusion and hospitalization variables are binary indicators. Facility elevation is measured in feet above sea level. The cut points between elevation quintiles are 73, 260, 599, and 870 feet above sea level.

D. Potential Anticipatory Responses

Due to frictions in changing clinical practices, we may expect them to change gradually and in anticipation of the PPS. Indeed, in Figure 1 we see that EPO doses began to decrease in mid-2010, prior to the PPS's start in January 2011 and shortly after the announcement of the final PPS rule in mid-August 2010. In this appendix, we both quantify these anticipatory effects and show that our results are robust to including this period of anticipatory responses by providers in the post-PPS period.

To identify and quantify this possible anticipation, we use the methods of Brot-Goldberg et al. (2017). First, we estimate

$$Y_{ijt} = \beta_0 + \beta_1 t + X_t \Gamma + \varepsilon_{ijt}, \tag{12}$$

where Y_{ijt} is the EPO dose for patient i at facility j in month t and X_t is a series of month-of-year fixed effects. We estimate this equation using only data from January 2005 through December 2009 and then use the estimated coefficients to calculate the predicted mean level of EPO for each month in 2010 and 2011. From the predicted and observed values in Table A6, we find that the first month in which the realized mean EPO dose is below the predicted level is October 2010, and that this drop continues to grow through 2011.

We corroborate our finding that the anticipatory response began in October 2010 by using a falsification test from Baicker and Svoronos (2019). To do so, we construct a test statistic from a series of Wald tests, testing each month in our data as a potential structural break in the time series of mean monthly EPO doses. From this, October 2010 returns the highest Wald statistic, 267, suggesting it is the most likely month of a structural break in the trend in EPO doses, which would indicate an anticipation of the PPS by providers.

In light of a possible anticipatory response, we consider the robustness of our main findings to this anticipation. In particular, we recreate the tables and figures presented in the main text while treating the start date of the PPS as October 2010 rather than the actual start date of January 2011. In this way, we treat the period during which facilities were modifying their behavior in anticipation of the PPS as part of the treatment period. Tables A7–A8 and Figure A4 recreate our main results and show that they are robust to this alternative definition of the PPS period.

Table A6
Difference in EPO Relative to Trend

	Actual	Predicted	Difference
2010			
January	58.95	56.19	2.76
February	55.81	52.28	3.53
March	63.36	57.90	5.46
April	59.39	55.96	3.43
May	58.64	58.08	0.56
June	59.06	56.60	2.46
July	59.63	57.64	1.99
August	57.76	57.76	0.00
September	55.77	55.77	0.00
October	53.57	57.61	-4.04
November	51.85	55.03	-3.17
December	50.80	56.94	-6.14
2011			
January	49.98	54.64	-4.66
February	45.90	50.72	-4.82
March	50.77	56.34	-5.57
April	48.88	54.41	-5.52
May	48.36	56.52	-8.16
June	47.80	55.04	-7.25
July	46.74	56.09	-9.35
August	42.97	56.20	-13.24
September	38.66	54.21	-15.55
October	39.01	56.05	-17.03
November	38.68	53.47	-14.79
December	38.65	55.39	-16.74

Notes: Predicted values from OLS estimate of (12). Dependent variable is monthly EPO dose. EPO doses are censored at the 99th percentile and measured in 1000 IUs. An observation is a patient-month. Estimation sample consists of observations from January 2005 to December 2009 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Sample presented in table consist of analogous observations from January 2010 to December 2011.

Table A7
The Effect of EPO on Anemia

	HGB		Transfusion	
	(1) OLS	(2) IV	(3) OLS	(4) IV
EPO	-0.00283***	0.0164***	0.000125***	-0.000570***
	(0.0000248)	(0.00460)	(0.00000250)	(0.000147)
Year-Month FE	1	1	1	1
Pat/Fac Controls	1	1	1	1
Facility FE	1	1	1	1
Dep. Var. Mean	11.17	11.17	0.0279	0.0279
Observations	8056164	8056164	9979284	9979284
First-Stage F-statistic		37.44		55.28

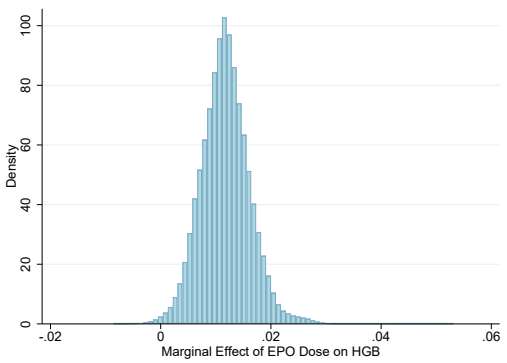
Notes: OLS and IV estimates from (7). Dependent variable in columns (1)–(2) is hemoglobin. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. Dependent variables in columns (3)–(4) is a binary outcome variable for receiving a blood transfusion. EPO doses are censored at the 99th percentile and measured in 1000 IUs. An observation is a patient-month. Sample consists of observations from October 2008 to September 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. Standard errors clustered by facility are in parentheses. $^{+}$, *, **, and *** indicate significance at the 10%, 5%, 1%, and 0.1% level, respectively.

Table A8
The Effect of EPO on Hospitalizations and Mortality

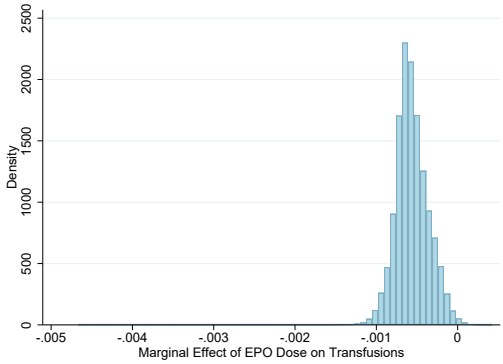
	Hosp., Any Cause		Hosp., Cardiac Event		Hosp., Septicemia		Mortality	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
EPO	0.000147***	0.0000722	0.0000146***	0.000119	-0.000000783	0.0000270	-0.000112***	0.000145*
	(0.00000343)	(0.000238)	(0.00000119)	(0.0000961)	(0.000000587)	(0.0000526)	(0.000000871)	(0.0000649)
Year-Month FE	1	1	1	1	1	1	1	1
Pat/Fac Controls	1	1	1	1	1	1	1	1
Facility FE	1	1	1	1	1	1	1	1
Dep. Var. Mean	0.139	0.139	0.0274	0.0274	0.00930	0.00930	0.0159	0.0159
Observations	9979284	9979284	9979284	9979284	9979284	9979284	9979284	9979284
First-Stage F-statistic		55.28		55.28		55.28		55.28

Notes: OLS and IV estimates from (7). Dependent variables are binary outcomes. EPO doses are censored at the 99th percentile and measured in 1000 IUs. An observation is a patient-month. Sample consists of observations from October 2008 to September 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. Standard errors clustered by facility are in parentheses. $^{+}$, *, **, and *** indicate significance at the 10%, 5%, 1%, and 0.1% level, respectively.

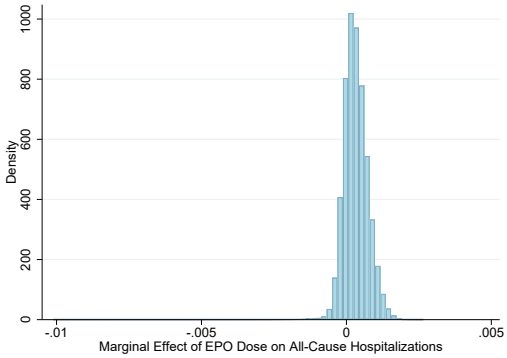
Figure A4
Histogram of Estimated Marginal Effects of EPO



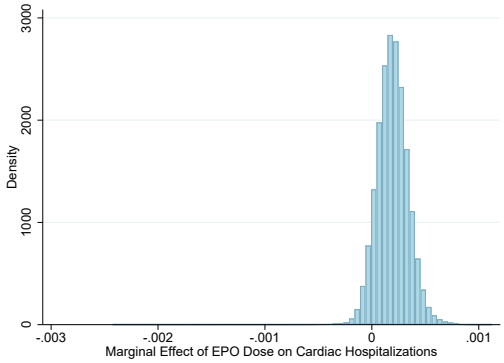
(a) HGB



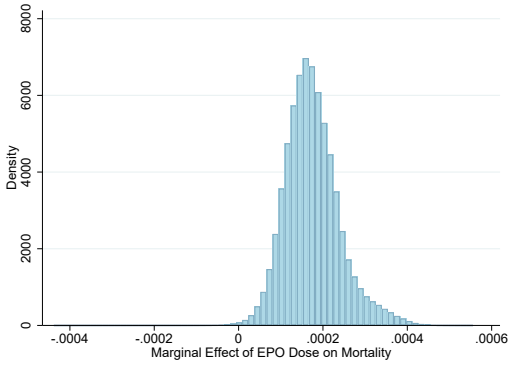
(b) Transfusions



(c) All-Cause Hospitalization



(d) Cardiac Hospitalization



(e) Mortality

Notes: Estimated marginal effects come from IV estimates of (10). An observation is a patient-month. Sample consists of observations from October 2008 to September 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter.

E. Investigating Identifying Assumptions

E.1. Other Drugs

In addition to EPO, intravenous iron and vitamin D are common classes of injectable drugs administered to dialysis patients. Like EPO, these were separately billable prior to 2011, but were then bundled together with dialysis in the payment reform. Unlike EPO, these drugs were not the subject of any changes in clinical guidelines, such as the black box warning for EPO issued by the FDA in mid-2011.

Any change in providers' use of these drugs in response to the payment reform may violate the exclusion restriction for identifying the marginal effect of EPO on health outcomes. To address this, we present an alternative approach in which we account for intravenous iron in addition to EPO, although we exclude vitamin D because it was not used to treat anemia. Table A9 presents the summary statistics with information on the use of these other injectable drugs, which are used much less often than EPO.

Table A9
Summary Statistics including the Use of Other Drugs

	Mean	Std. Dev.
Resource Use		
EPO Dose (1000 IUs)	48.50	64.11
Receives Any EPO	0.755	0.430
IV Iron Dose (1000 IUs)	0.20	0.26
Receives Any Iron	0.571	0.495
Vitamin D Dose (1000 IUs)	0.03	0.06
Receives Any Vitamin D	0.659	0.474
Receives Any Cinacalcet	0.099	0.299
Dialysis Sessions	12.08	9.90
Unique Patients	461,477	
Patient-Months	10,077,289	

Notes: An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Injectable iron drugs include Ferrlecit, Venofer, Ferumoxytol, and Iron Dextran. Injectable vitamin D drugs include Calcitriol, Doxercalciferol, and Paricalcitol.

We re-estimate our main specification from Section 6.1 using a combined measure of intravenous

iron and EPO as our instrumented variable. Specifically, in each month we calculate each patient's Z-score for EPO based on the mean and standard deviation of EPO in our entire sample as well as a Z-score for intravenous iron. We average those together for a combined total anemia drug dose Z-score, which captures each patient's position in the distribution of total anemia drug use. The results are presented in Table A10 and are very similar to our baseline results, demonstrating their robustness.

Table A10
Combined Injectable Anemia Drugs and Outcomes

	(1) HGB	(2) Transfusion	(3) Mortality	(4) Hosp., Any Cause	(5) Hosp., Cardiac Event	(6) Hosp., Septicemia
Combined Injectibles Z-score	1.594***	-0.0471***	0.0103^{+}	0.0163	0.0148^{+}	0.00287
	(0.386)	(0.0127)	(0.00533)	(0.0206)	(0.00796)	(0.00441)
Year-Month FE	1	1	1	1	1	1
Pat/Fac Controls	1	1	1	1	1	1
Facility FE	1	1	1	1	1	1
Dep. Var. Mean	11.12	0.0282	0.0157	0.138	0.0271	0.00939
Observations	8181736	10077264	10077264	10077264	10077264	10077264
First-Stage F-statistic	33.42	38.24	38.24	38.24	38.24	38.24

Notes: IV estimates from (7). Dependent variable in column (1) is hemoglobin. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. Dependent variables in columns (2)–(6) are binary outcomes. Combined injectables Z-score is the mean of the patient-month's Z-scores for EPO use and IV iron use. Injectable iron drugs include Ferrlecit, Venofer, Ferumoxytol, and Iron Dextran. An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. Standard errors clustered by facility are in parentheses. $^{+}$, *, **, and *** indicate significance at the 10%, 5%, 1%, and 0.1% level, respectively.

E.2. Dialysis Modality

As discussed by (Zhang et al., 2017), a change in relative profitability after the PPS that favored peritoneal dialysis may have led providers to shift patients away from in-center hemodialysis. Like the use of other anemia drugs, this change may violate the exclusion restriction for identifying the marginal effect of EPO on health outcomes. In Table A11, we show that neither the share of patients receiving in-center hemodialysis nor the share receiving peritoneal dialysis changed differentially by elevation after the PPS, further supporting our identification strategy.

Table A11
Differential Change by Elevation for Dialysis Modality

	(1) Dialysis Sessions	(2) In-Center Hemodialysis	(3) Peritoneal Dialysis	(4) Good URR
Elevation × PPS	-0.00000473	-0.000000666	0.000000432	0.000000980
	(0.00000560)	(0.000000854)	(0.000000802)	(0.000000696)
Year-Month FE	1	1	1	1
Pat/Fac Controls	1	1	1	1
Facility FE	1	1	1	1
R-squared	0.00617	0.281	0.264	0.0923
Dep. Var. Mean	12.08	0.909	0.0702	0.933
Observations	8869420	10267613	10267613	8560825

Notes: OLS estimates from (8). Dependent variable in column (1) is monthly number of dialysis sessions, in column (2) is an indicator for receiving in-center hemodialysis treatment, in column (3) is an indicator for receiving peritoneal dialysis treatment, and in column (4) is an indicator for having a urea reduction ratio above 0.85. PPS is an indicator variable for January 2011 or later. Facility elevation is measured in feet above sea level. An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for ESRD patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. In columns (1) and (4), the sample is further limited to in-center hemodialysis patients. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. Standard errors clustered by facility are in parentheses. $^{+}$, *, **, and *** indicate significance at the 10%, 5%, 1%, and 0.1% level, respectively.

E.3. Exclusion of Elevation

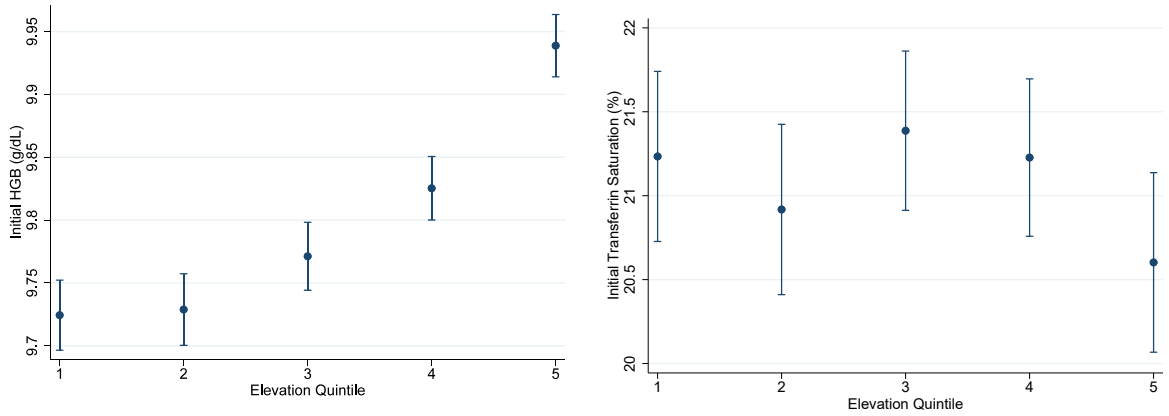
We allow the baseline observable health outcomes that patients would attain in the absence of EPO to differ by elevation, but we assume that the causal effect of EPO on observable health outcomes does not systematically vary with it, a conclusion supported by the medical literature (see the discussion in Section 2). In this section, we provide new evidence from our data that supports the findings in that literature. We then provide corroborating evidence more generally of a constant marginal effect of EPO by elevation.

As discussed in Section 2.1, our assumption that the causal effect of EPO on observable health outcomes does not systematically differ by elevation is consistent with existing evidence from the medical literature. To support this conclusion, we first provide evidence that patients' baseline HGB levels—those they would attain without supplemental EPO—vary by elevation. This is the reason that patients at different elevations require different amounts of EPO. We demonstrate that in our sample, HGB levels observed before patients begin supplemental EPO are positively correlated with elevation. Panel (a) of Figure A5 and the first three columns of Table A12 demonstrate this—the HGB levels patients had at their first dialysis treatment before receiving EPO increase with elevation. The clear positive relationship between pre-EPO HGB levels and elevation also shows the importance of allowing baseline health outcome levels to differ by elevation.

To more directly explore the assumption that the effect of EPO does not systematically vary by elevation, we explore how transferrin saturation rates (TSAT) differ by elevation, again focusing on measures prior to EPO treatment. As discussed in Section 2.1, iron availability is the most likely channel through which elevation could alter the returns from EPO—more iron could result in more productive EPO. Transferrin saturation is the best measure in our data of iron availability in the bloodstream at the time of treatment. Transferrin is a protein that transports iron through the bloodstream and transferrin saturation measures the share of transferrin that is being used to transport iron (Nakhoul and Simon, 2016). While elevation-induced hypoxia stimulates increased iron release into the bloodstream, there is a simultaneous increase in natural erythropoietin and some of the iron is exhausted in the resulting red blood cell production. What matters for the returns of supplemental EPO is the remaining iron available when EPO is administered, as measured by TSAT. Panel (b) of Figure A5 and columns (4)–(6) of Table A12 show that TSAT among patients beginning dialysis is remarkably constant across different

elevations, replicating the findings of Sibbel et al. (2017) in our sample. Taken together, this evidence suggests that the amount of EPO patients require depends on elevation because their baseline HGB levels vary with elevation, and not because the marginal impact of EPO varies with elevation.

Figure A5
HGB Differences by Elevation



(a) Initial HGB

(b) Initial Transferrin Saturation Value

Notes: An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. The sample is further limited to the first month the patient receives dialysis. Panel (a) gives the average HGB at the beginning of the month by elevation quintile. Panel (b) gives the average transferrin saturation rate at the beginning of the month for first-time dialysis patients in 2012, the first year in which this data were available. Facility elevation is measured in feet above sea level. The cut points between elevation quintiles are 73, 260, 599, and 870 feet above sea level. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. 95% confidence intervals are shown. Standard errors are clustered by facility.

We also appeal to the literature on treatment effect heterogeneity to further support the conclusion that the effect of EPO does not vary with elevation. In general, unobserved treatment effect heterogeneity, when correlated with the instrument, can violate both the independence assumption and exclusion restriction. Ignoring potential heterogeneity in the treatment effect by elevation could result in a biased estimator. When treatment effect heterogeneity is a concern, a natural response is to test whether treatment effects vary by observable covariates (Mogstad and Torgovitsky, 2024; Goldsmith-Pinkham et al., 2020), as we do in equation (10). In this case, however, enriching the model to allow for treatment effect heterogeneity by elevation poses a challenge because it absorbs some of the explanatory power of the instrument, which itself depends on elevation. In column (1) of Table A13 below, we present estimates from a version of equation (10), where we interact both the instrument and EPO with indicators for

Table A12
Initial Differences by Elevation

	(1) Initial HGB	(2) Initial HGB	(3) Initial HGB	(4) Initial TSAT	(5) Initial TSAT	(6) Initial TSAT
Elevation (1000 ft.)	0.0909***	0.0111**	0.0115**	-0.0974	0.0543	0.0694
	(0.00760)	(0.00357)	(0.00354)	(0.147)	(0.224)	(0.224)
Year-Month FE	0	0	1	0	0	1
Pat/Fac Controls	0	1	1	0	1	1
Facility FE	0	0	0	0	0	0
Dep. Var. Mean	9.801	9.801	9.801	21.07	21.07	21.07
Observations	149363	149363	149363	11950	11950	11950

Notes: Estimates from a regression of the dependent variable listed in each column on elevation. Dependent variable in columns (1)–(3) is HGB at the beginning of the month. Dependent variable in columns (4)–(6) is transferrin saturation. Facility elevation is measured in thousands of feet above sea level. An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. The sample is further limited to the first month the patient receives dialysis. For columns (4)–(6), the sample is limited to 2012. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include facility elevation, whether the facility is freestanding or hospital-based, and chain ownership status, as well as geographic information about the patient's residence, including state and average income in the ZIP code. Standard errors clustered by facility are in parentheses. $^{+}$, *, **, and *** indicate significance at the 10%, 5%, 1%, and 0.1% level, respectively.

the quintile of facility elevation. The similarity of the coefficient estimates across quintiles is strongly suggestive of homogeneous effects, with a joint F-test of the equality of the coefficients having a P-value of 0.679. That these estimates are larger than our main results in Table 5 is likely due to the weakened first stage, as demonstrated by the reported first-stage F-statistic.

In column (2) of Table A13, we omit the time fixed effects and instead add month-of-year fixed effects while using PPS—rather than PPS interacted with elevation—as the IV. Of course, omitting time fixed effects risks conflating other time trends with the changes in EPO and may result in biased and inconsistent point estimates. However, under the assumption that any potential time series confounders do not vary by elevation (i.e., that $Cov(PPS_t, \varepsilon_{ijt}|Elevation_j)$ is constant across elevation) we can identify *heterogeneity* in the point estimates by elevation even if the point estimates themselves are inconsistent.[27] That is, each coefficient should be biased *in the same way* under the assumption that there are no differences in time trends across elevation.[28] Here, we have a much stronger first stage and very precise estimates, and again, we find no evidence of heterogeneity by elevation, failing to reject the null hypothesis that the coefficients are equal. Thus, across both specifications we find no evidence of

[27]Note that this is a slightly stronger assumption than that necessary for the IV we use in the main text to be valid. In particular, we now need the instrument to be exogenous at each level of elevation, whereas before, we only needed the instrument to be exogenous *on average*.

[28]Insofar as the strength of the first stage differs by elevation, the bias will be the same magnitude if there is no confounding or if the confounder scales with the strength of the first stage.

meaningful heterogeneity in the marginal effect of EPO on HGB by elevation.

Table A13
Tests of Homogeneity in Effect of EPO by Elevation

	HGB	
	(1) Elevation×PPS	(2) PPS
EPO - First Quintile	0.0335*	0.0209***
	(0.0168)	(0.000693)
EPO - Second Quintile	0.0345*	0.0217***
	(0.0172)	(0.000743)
EPO - Third Quintile	0.0342+	0.0211***
	(0.0177)	(0.000793)
EPO - Fourth Quintile	0.0378+	0.0233***
	(0.0197)	(0.00104)
EPO - Fifth Quintile	0.0361+	0.0213***
	(0.0195)	(0.000794)
Year-Month FE	1	0
Month FE	1	1
Pat/Fac Controls	1	1
Facility FE	1	1
Dep. Var. Mean	11.12	11.12
Observations	8181736	8181736
First-Stage F-statistic	0.964	258.6
Joint Equality P-Value	0.679	0.259

Notes: IV estimates from (7) with interactions between EPO and quintile of elevation. Dependent variable is hemoglobin. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. IV in column (1) is elevation interacted with PPS. IV in column (2) is PPS. EPO doses are censored at the 99th percentile and measured in 1000 IUs. An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Further controls include calendar month fixed effects. Column (1) includes month-year fixed effects. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. Standard errors clustered by facility are in parentheses. +, *, **, and *** indicate significance at the 10%, 5%, 1%, and 0.1% level, respectively.

E.4. Characterizing Compliers

Following Angrist and Pischke (2009), we can conceptualize our 2SLS estimates as the weighted average of the potentially nonlinear causal effects of EPO on the outcome, where the weighting comes from the number of compliers at each level of EPO, which relies on the usual assumptions about the validity of the instrument (exclusion/independence, existence of the first stage, and monotonicity). Given those assumptions hold, the size of the complier group can be written as

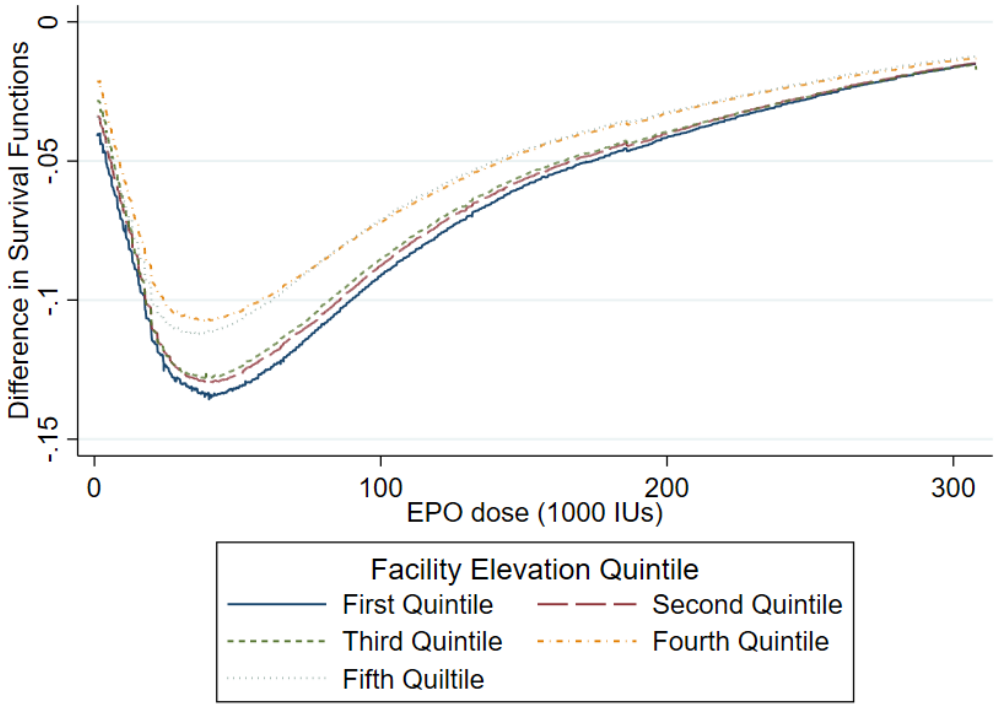
$$P[e_{1i} \leq e < e_{0i}] = P[e_{1i} \leq e] - P[e_{0i} < e],$$

where e_{1i} is the EPO given to patient i when the instrument is "switched on" and e_{0i} is the amount given when the instrument is "switched off." Thus, the measure of compliers at some EPO level e is the measure of patients who are induced to receive a dose less than e by the instrument and would not otherwise. In other words, the weighting function compares the CDFs of the endogenous variable across different values of the instrument.

Because our instrument is continuous, it is difficult to analytically recover the exact weights placed on observations with different characteristics. However, Figure A6 gives a sense of how these weights are distributed. The figure plots the differences in the survival functions for EPO doses in the pre-PPS period and the post-PPS period for the five elevation quintiles, where the reference groups are the pre-PPS quintiles. In other words, each line is the difference for a given elevation quintile between the probability that a patient receives or exceeds the dose indicated on the x-axis *before* the policy change and the probability that the patient receives or exceeds that dose *after* the policy change.

Figure A6 provides a number of insights. First, compliance is near universal. Across the entire range of EPO doses and for all elevations, patients tend to receive lower doses after the policy change, with larger differences for the low elevations (quintiles 1-3) than the high elevations (quintiles 4-5). Second, we can observe where the weighting from our instrument comes from: the differences between the high and low elevation differences plotted. There are compliers at nearly all levels of EPO, particularly at moderate levels of EPO common among patients. This indicates that rather than exploiting variation only for very specific amounts of EPO, our analysis uses variation across the distribution. In summary, this analysis suggests that the local average treatment effect we recover is indeed informative of the

Figure A6
Difference in EPO Survival Function by Elevation



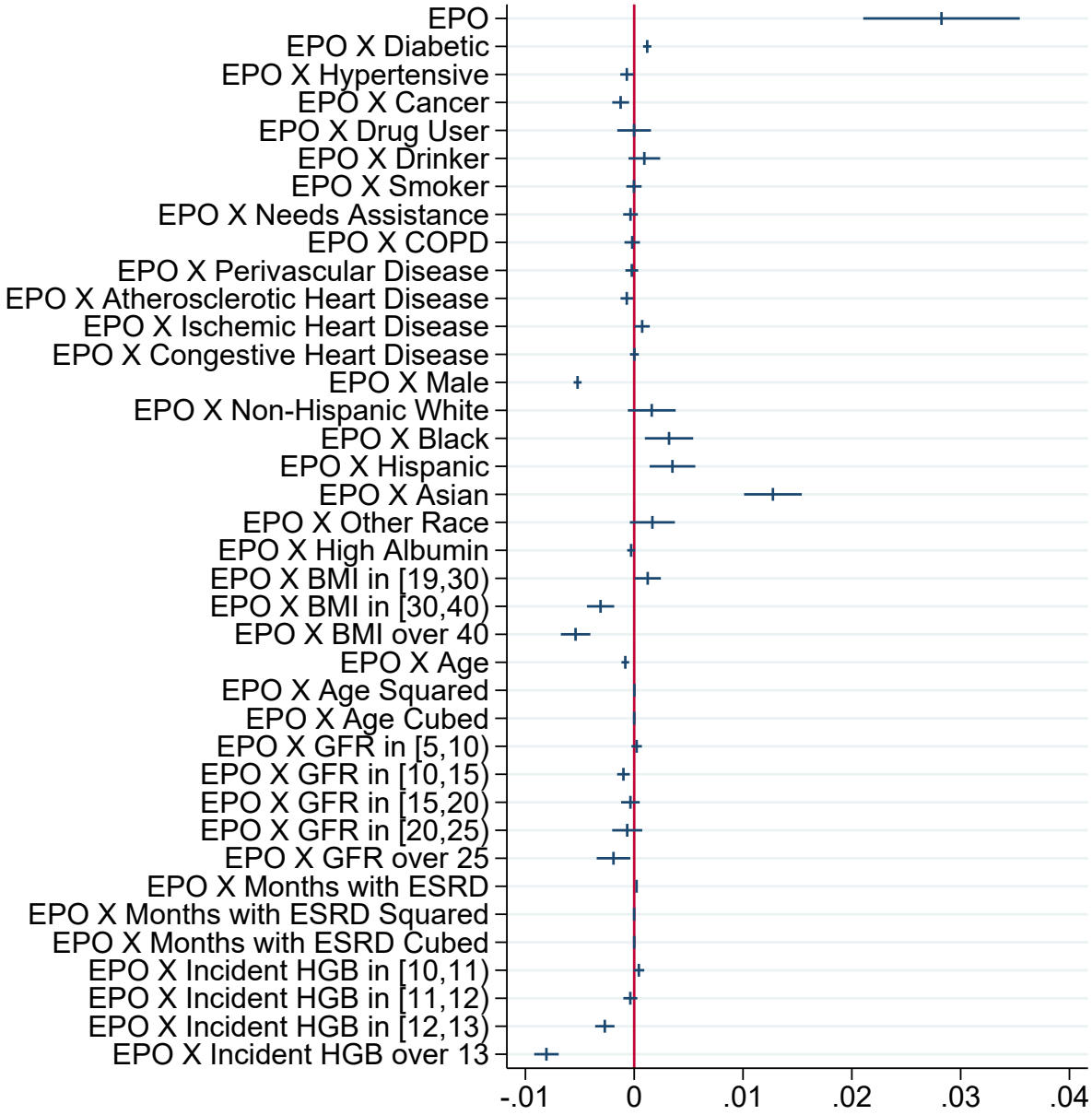
Notes: Figure plots the difference in the survival function for EPO doses between observations from 2009–2010 and 2011–2012. An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Facility elevation is measured in feet above sea level. The cut points between elevation quintiles are 73, 260, 599, and 870 feet above sea level. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter.

broader, policy-relevant effect.

F. Additional Parameter Estimates

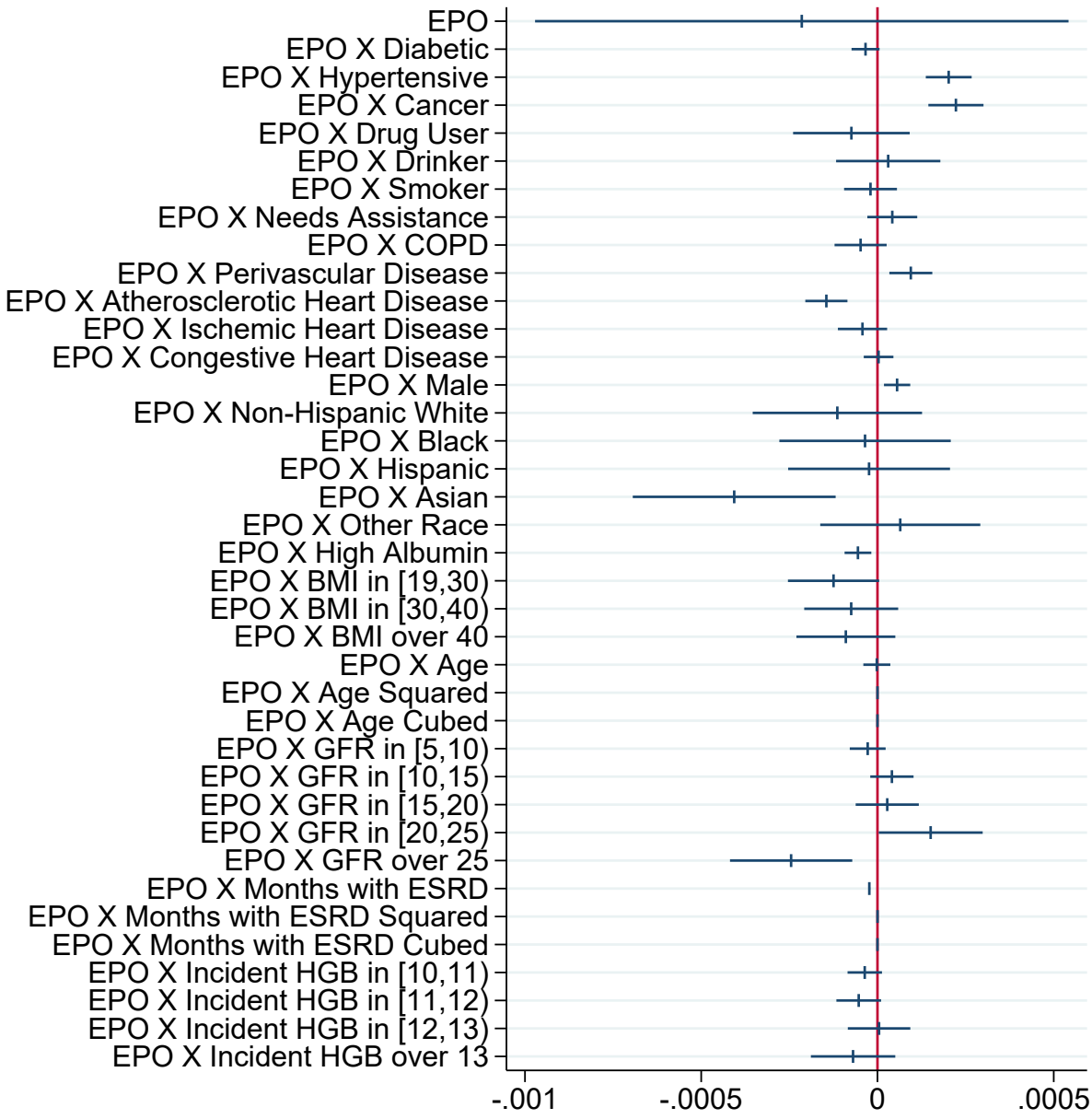
In this appendix, we report estimates of parameter values not reported in the main text. Figures A7–A11 report estimates of $\alpha_{c,X}^{IV}$, which are also $\alpha_{c,x}$ in structural model (11). Figures A12–A16 report the CDF of the implied estimated marginal effects along with 90%, 95%, and 99% confidence intervals from estimation on 200 bootstrap samples.

Figure A7
Coefficient Estimates of Heterogeneity in Responsiveness of HGB to EPO



Notes: IV estimates from (10). An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. Horizontal bands give 95% confidence intervals. Standard errors are clustered at the facility level.

Figure A8
Coefficient Estimates of Heterogeneity in Responsiveness of Transfusions to EPO

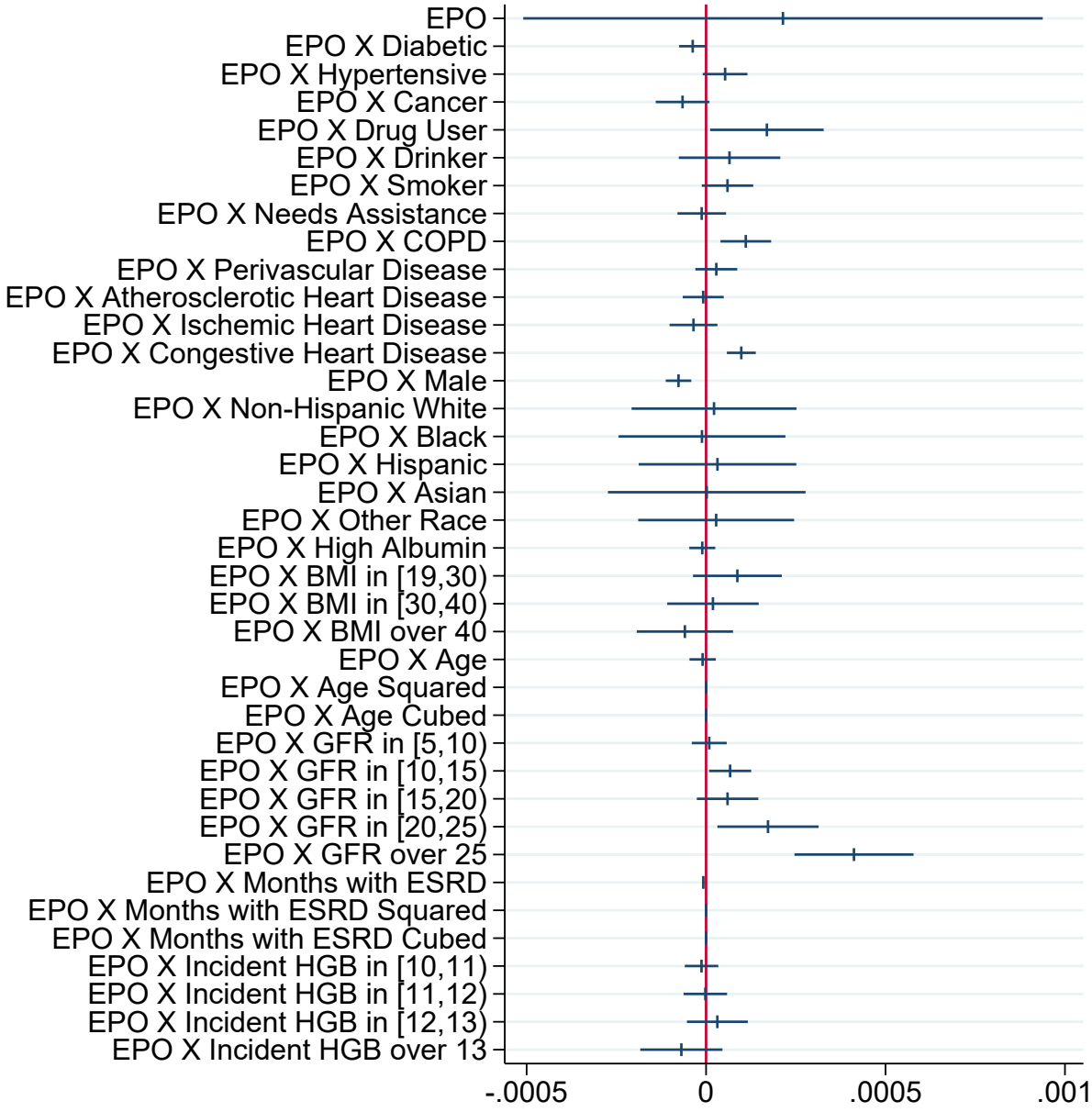


Notes: IV estimates from (10). An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Horizontal bands give 95% confidence intervals. Standard errors are clustered at the facility level.

Figure A9
Coefficient Estimates of Heterogeneity in Responsiveness of All-Cause Hospitalizations to EPO

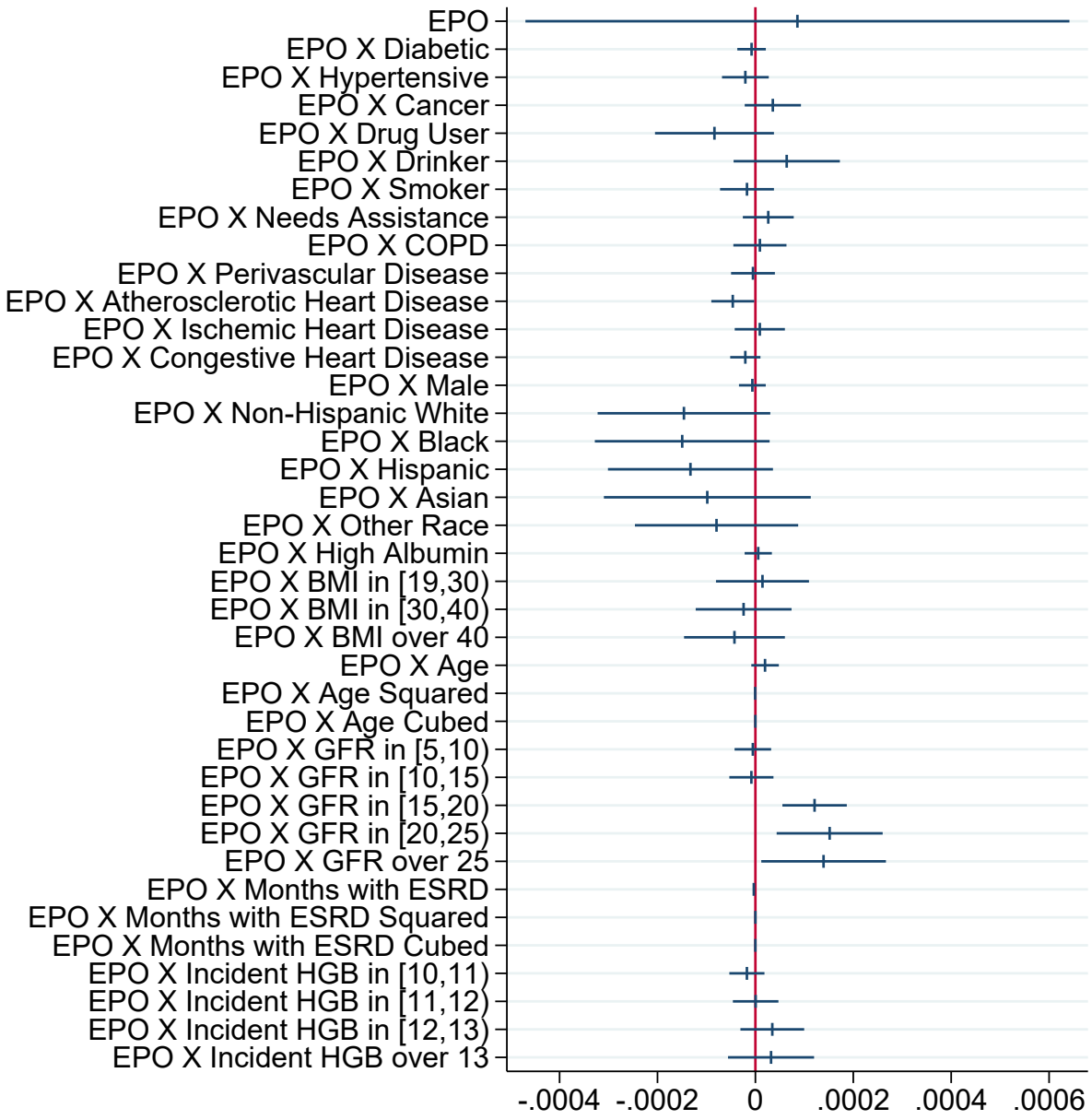
Notes: IV estimates from (10). An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Horizontal bands give 95% confidence intervals. Standard errors are clustered at the facility level.

Figure A10
Coefficient Estimates of Heterogeneity in Responsiveness of Cardiac Hospitalizations to EPO



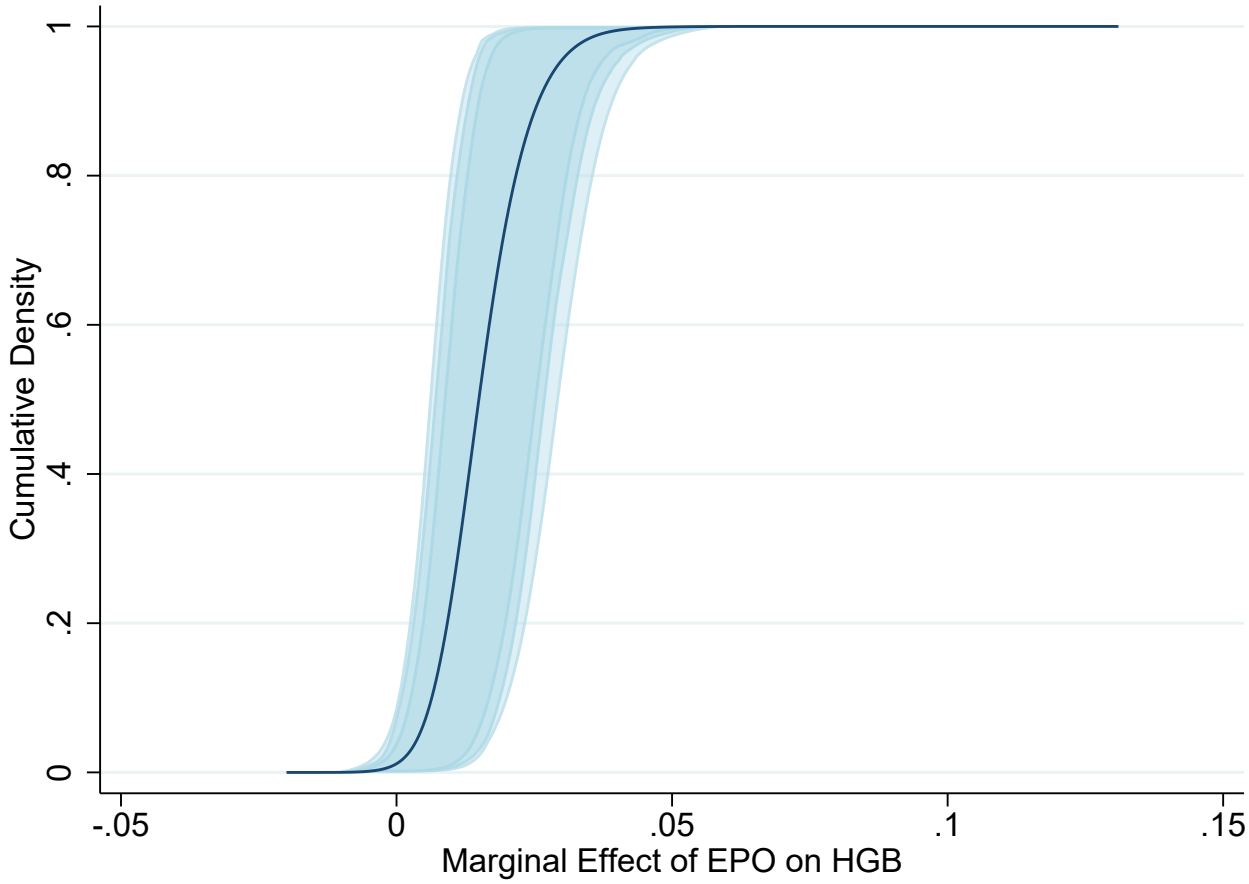
Notes: IV estimates from (10). An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Horizontal bands give 95% confidence intervals. Standard errors are clustered at the facility level.

Figure A11
Coefficient Estimates of Heterogeneity in Responsiveness of Mortality to EPO



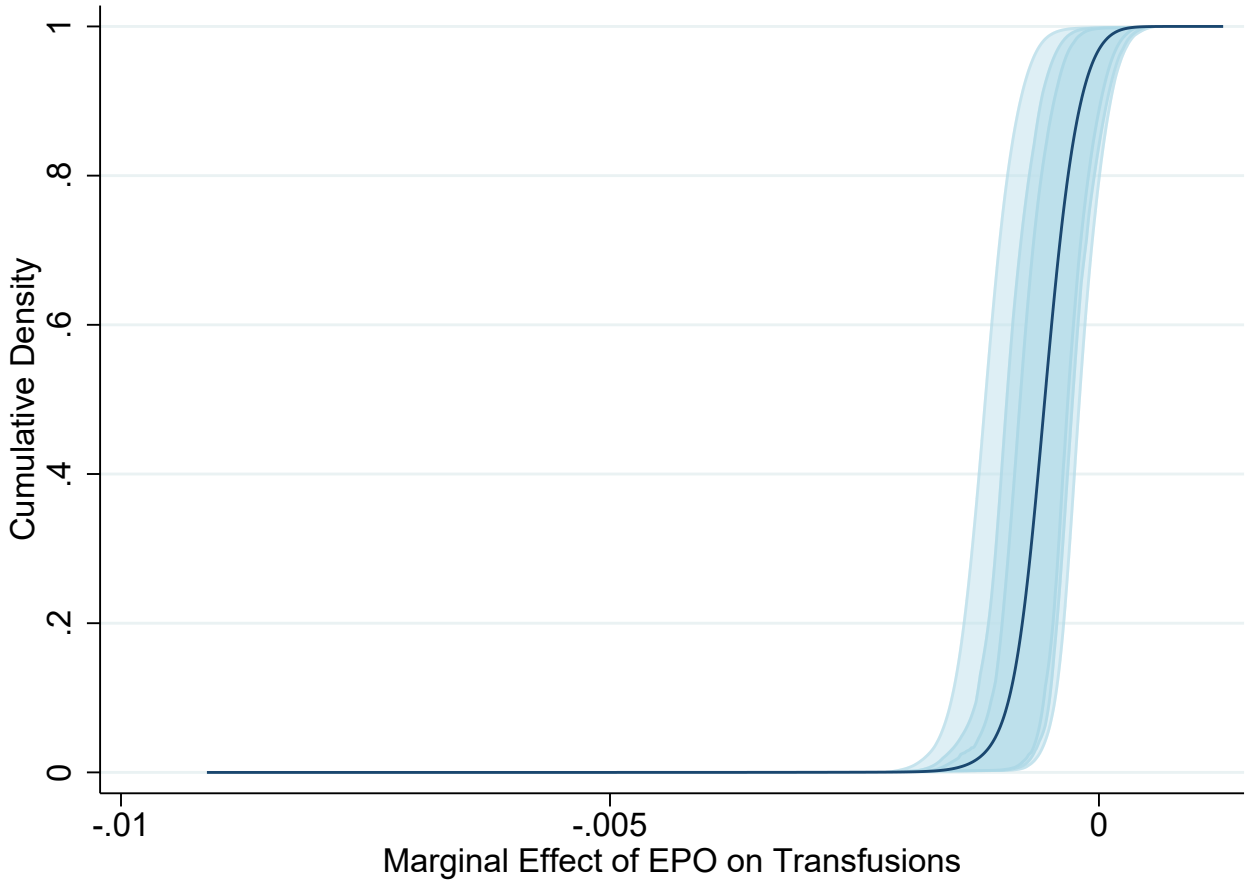
Notes: IV estimates from (10). An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Horizontal bands give 95% confidence intervals. Standard errors are clustered at the facility level.

Figure A12
Estimated CDF of Responsiveness of HGB to EPO



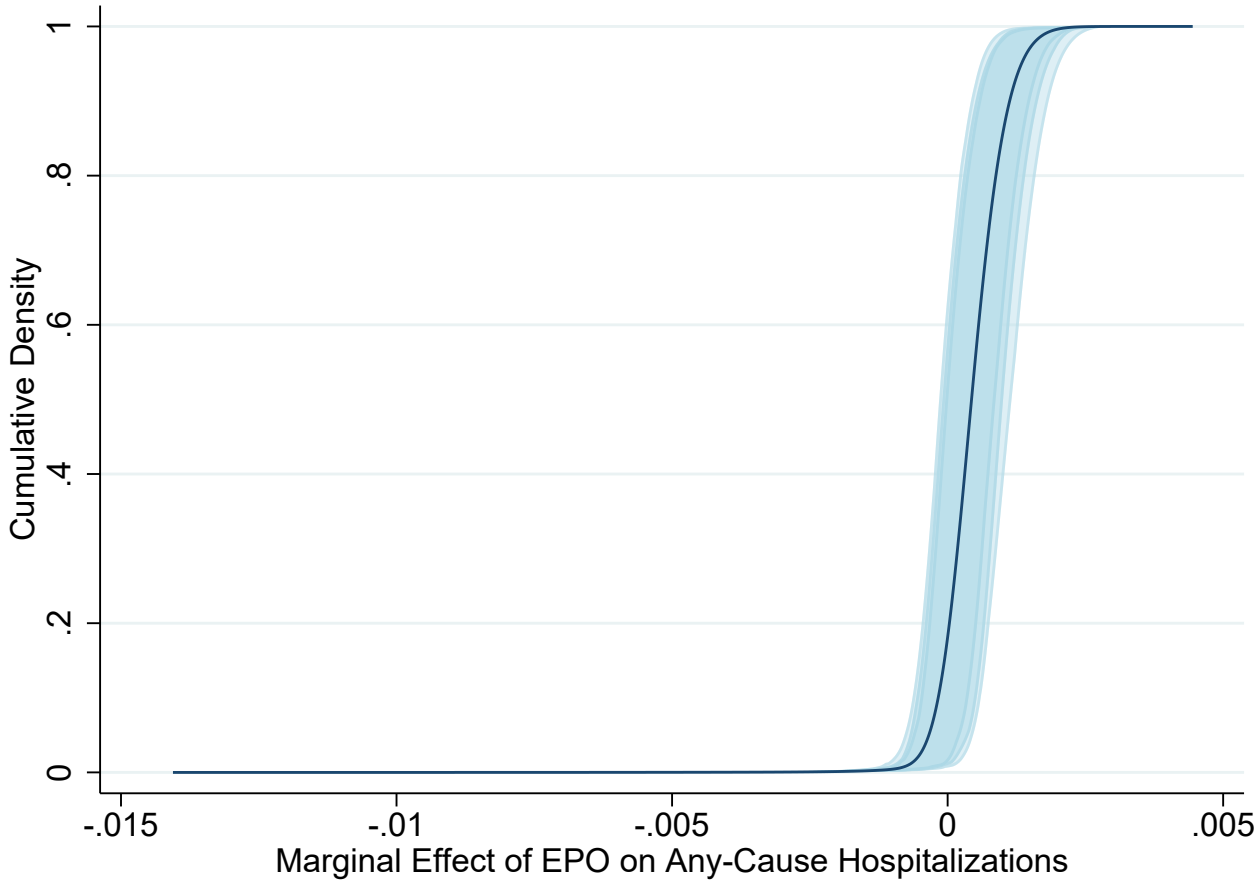
Notes: Distribution of estimated marginal effect of EPO on HGB from IV estimates of (10). An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. Darkest blue region gives the 90% confidence interval, while lighter regions give the 95% and 99% confidence intervals. Confidence intervals are calculated from estimation on 200 bootstrap samples, sampled at the facility level.

Figure A13
Estimated CDF of Responsiveness of Transfusions to EPO



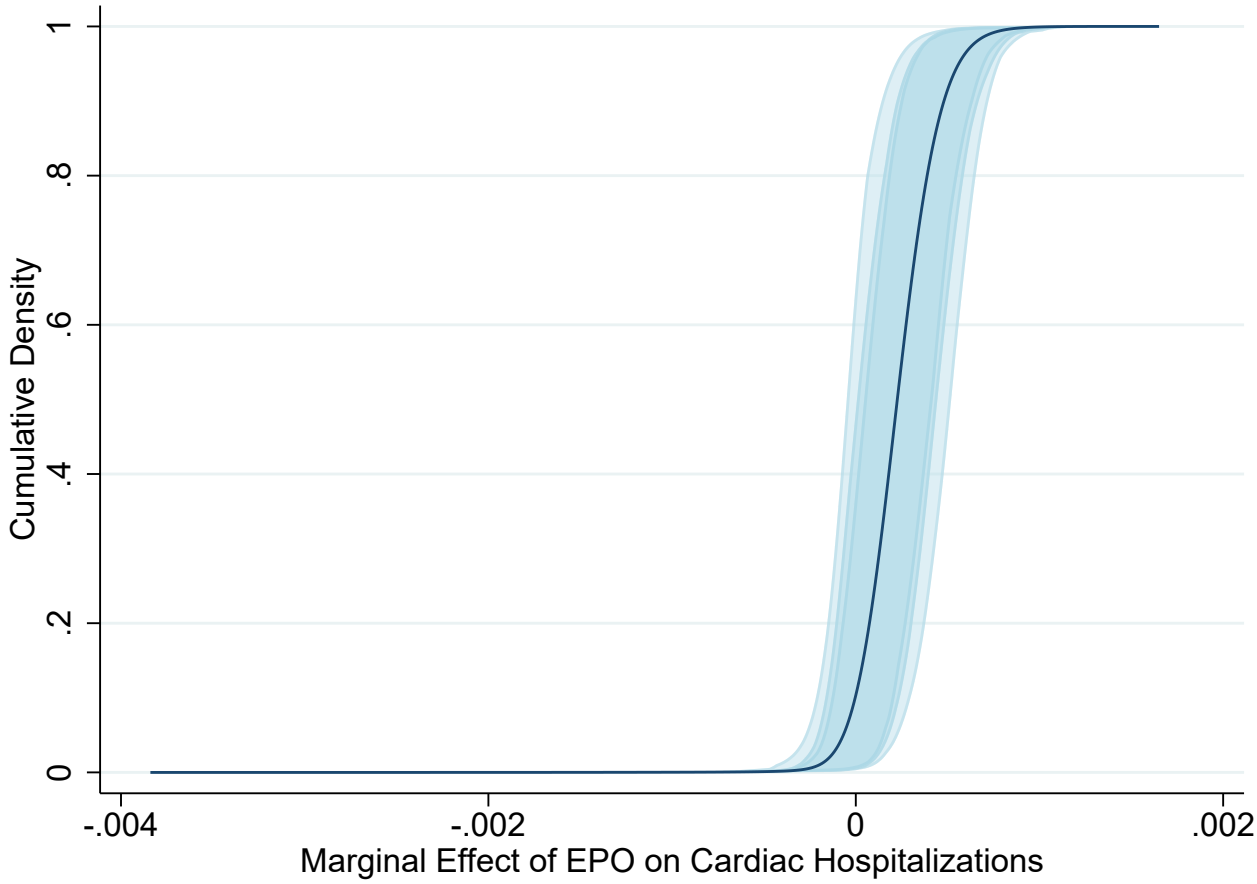
Notes: Distribution of estimated marginal effect of EPO on transfusions from IV estimates of (10). An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Darkest blue region gives the 90% confidence interval, while lighter regions give the 95% and 99% confidence intervals. Confidence intervals are calculated from estimation on 200 bootstrap samples, sampled at the facility level.

Figure A14
Estimated CDF of Responsiveness of All-Cause Hospitalizations to EPO



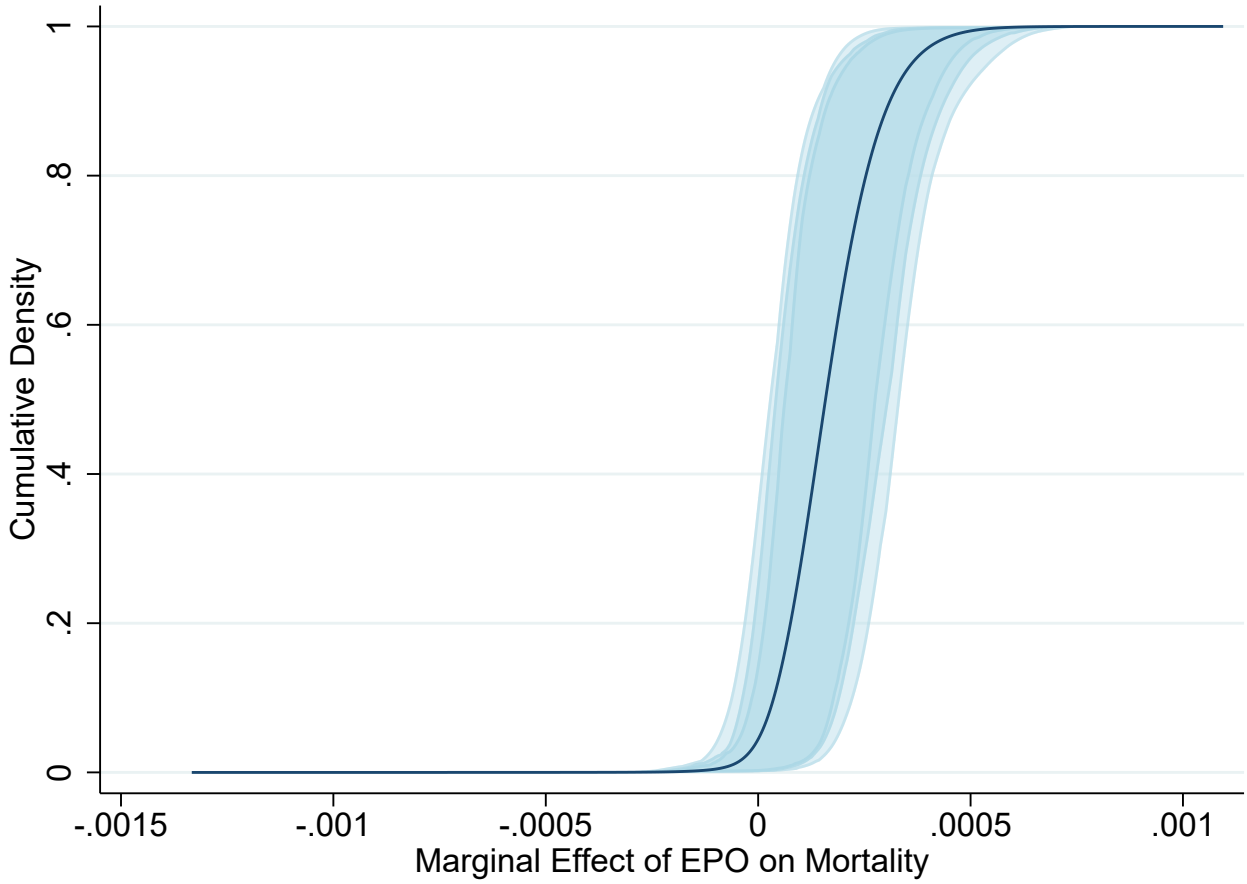
Notes: Distribution of estimated marginal effect of EPO on hospitalizations from IV estimates of (10). An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Darkest blue region gives the 90% confidence interval, while lighter regions give the 95% and 99% confidence intervals. Confidence intervals are calculated from estimation on 200 bootstrap samples, sampled at the facility level.

Figure A15
Estimated CDF of Responsiveness of Cardiac Hospitalizations to EPO



Notes: Distribution of estimated marginal effect of EPO on cardiac hospitalizations from IV estimates of (10). An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Darkest blue region gives the 90% confidence interval, while lighter regions give the 95% and 99% confidence intervals. Confidence intervals are calculated from estimation on 200 bootstrap samples, sampled at the facility level.

Figure A16
Estimated CDF of Responsiveness of Mortality to EPO



Notes: Distribution of estimated marginal effect of EPO on mortality from IV estimates of (10). An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. Patient controls include dummy variables for incident comorbidities and characteristics reported on medical evidence forms, including diabetes, hypertension, BMI bin, GFR bin, HGB bin, high albumin, cancer, drug use, alcoholism, smoking behavior, necessity of assistance, COPD, ASHD, PVD, ischemic heart disease, and congestive heart disease, along with patient race, gender, and cubic functions of age and dialysis tenure. Facility controls include whether the facility is freestanding or hospital-based, and chain ownership, as well as facility fixed effects and geographic information about the patient's residence, including state and average income in the ZIP code. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Darkest blue region gives the 90% confidence interval, while lighter regions give the 95% and 99% confidence intervals. Confidence intervals are calculated from estimation on 200 bootstrap samples, sampled at the facility level.

G. Additional Structural Model results

In this appendix, we report additional results from estimation of our structural model that we reference in the main text. Table A14 reports the estimated mean and standard deviation of costs by chain ownership. Note that the mean costs for each chain are calibrated to match the mean EPO acquisition costs reported in the 2010 cost report data. Table A16 reports the changes in counterfactual values of EPO use and observed health outcomes relative to baseline in percentage terms. Note that we do not report the change in unobserved or total health because, while we are able to estimate level changes in these variables, we do not observe the baseline levels.

Table A14
Estimated EPO Acquisition Costs by Chain

	Mean	Standard Deviation	Obs.
DaVita	7.92	3.19	3192294
Fresenius	7.33	3.85	3471080
Independents	8.54	16.9	1980611
Other Chains	8.36	3.35	1433304
Overall	7.90	8.13	10077289

Notes: Estimated EPO acquisition costs by chain. Mean acquisition costs are calibrated to match 2010 cost report data and are reported in dollars per 1,000 IUs. An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later.

Table A15
Observed and Model-Predicted Moment Values

	Data	Model-Predicted
Mass in 25th to 75th Percentiles of EPO Distribution	0.489	0.529
Mass at 0 or Maximum EPO Dose	0.127	0.168
Mean EPO Dose	54.486	53.473
Mean EPO Dose for Observations with Baseline HGB between 10 and 11 g/dL	40.178	36.661
Mean EPO Dose for Observations with Baseline HGB over 11 g/dL	24.187	23.430
Covariance between EPO and Baseline HGB	-57.833	-71.132
Covariance between EPO and Financial Margin on EPO	53.505	52.612
Covariance between EPO and Marginal Effect on HGB	-5.060	-6.596
Covariance between EPO and Marginal Effect on Cardiac Hospitalizations	-3.344	-4.725
Covariance between EPO and Marginal Effect on Mortality	-2.910	-5.557

Notes: Observed and model-predicted moment values. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. The sample is further limited to patient-months with an estimated baseline HGB level between 5 and 15 and an estimated effect of EPO on HGB of greater than 0.005.

Table A16
Percent Change in Counterfactual Outcomes

	(1) Health-Maximizing EPO Dose	(2) Uniform EPO Dose	(3) Minimum HGB Target	(4) Chain Altruism Equal to Independent Altruism	(5) Lower Acquisition Cost of EPO
EPO Dose (%)	36.59	36.59	6.491	0.064	2.740
Hemoglobin (%)	1.889	2.349	0.294	-0.095	0.150
Transfusion Rate (%)	-19.112	-19.498	-16.911	-14.254	-15.436
All-Cause Hospitalizations (%)	4.054	6.335	5.369	3.766	4.699
Cardiac Hospitalization (%)	11.487	16.899	13.963	9.744	12.425
Mortality (%)	16.424	20.507	18.380	14.389	16.146

Notes: Change under counterfactual scenarios. The baseline for all columns is the simulated outcome under PPS. EPO doses are censored at the 99th percentile and measured in 1000 IUs. Counterfactual EPO doses are restricted to the observed support of doses. Hemoglobin is winsorized from below to 5 and from above to 20 and is measured in grams per deciliter. An observation is a patient-month. Sample consists of observations from January 2009 to December 2012 for in-center hemodialysis patients between the ages of 18 and 100 with Medicare as their primary payer for whom we observe all patient and facility controls used in the analyses in Section 5.1 and later. The sample is further limited to patient-months with an estimated baseline HGB level between 5 and 15 and an estimated effect of EPO on HGB of greater than 0.005.