Supply-side solutions targeting demand-side characteristics: causal effects of a chronic disease management program on adherence and health outcomes

SUPPLY-SIDE SOLUTIONS TARGETING DEMAND-SIDE CHARACTERISTICS: CAUSAL EFFECTS OF A CHRONIC DISEASE MANAGEMENT PROGRAM ON ADHERENCE AND HEALTH OUTCOMES

Luke Connelly, *Centre for the Business and Economics of Health, The University of Queensland, Brisbane, Australia; Dipartimento di Sociologia e Diritto dell'Economia, Università di Bologna, Bologna, Italy*

Gianluca Fiorentini, *Dipartimento di Scienze Economiche, Università di Bologna, Bologna, Italy*

Marica Iommi, *Scuola Superiore di Politiche per la Salute, Università di Bologna, Bologna, Italy*

ABSTRACT

We estimate the effects of a chronic disease management program (CDMP) which adapts various supply-side interventions to specific demand-side conditions (disease-staging) for patients with Chronic Kidney Disease (CKD). Using a unique dataset on the entire population of the Emilia-Romagna region of Italy with hospital-diagnosed CKD, we estimate the causal effects of the CDMP on adherence indicators and health outcomes. As CKD is a progressive disease with clearly-defined disease stages and a treatment regimen that can be titrated by disease severity, we calculate dynamic, severity-specific, indicators of adherence as well as several long-term health outcomes. Our empirical work produces statistically significant and sizeable causal effects on many adherence and health outcome indicators across all CKD patients. More interestingly, we show that the CDMP produces larger effects on patients with early-stage CKD, which is at odds with some of the literature on CDMP that advocates intensifying interventions for high-cost (or late-stage) patients. Our results suggest that it may be more efficient to target early-stage patients to slow the deterioration of their health capital. The results contribute to a small, recent literature in health economics that focuses on the marginal effectiveness of CDMPs after controlling either for supply- or demand-side sources of heterogeneity.

Keywords: Chronic kidney disease, casual effect, guidelines, adherence, survival.

JEL Codes: I11, I12, I18.

1 INTRODUCTION

Chronic diseases (CDs) account for a large proportion of health care expenditure in developed nations. In the United States, for example, some authors have suggested that 75% of total health care expenditures, 83% of Medicare expenditures, and 96% of Medicaid expenditures are attributable to chronic disease [1]. Furthermore, as life expectancy grows, the prevalence of CDs is also expected to rise across the developed world. In Europe, men's and women's life expectancy recently increased by 3.2 and 2.5 years, respectively, over the course of just one decade [2]. Life expectancies for men and women of age 65 are now 18 and 21 years, respectively, but it is estimated that only nine of those years is lived, on average, in good health [3]. Thus, according to the most recent Global Burden of Disease (GBD) study, chronic (or "non-communicable") diseases are now responsible for 97% of all deaths and 87% of Years Lived with Disability in Europe [4].

Accordingly, secondary prevention or "CD management" is a high priority in many countries. Specifically, the implementation of initiatives that are designed to slow disease progression by improving patients' adherence to clinical recommendations (e.g., in respect of medication use, laboratory tests and clinical visits) and the coordination of patients' care (e.g., as between generalist and specialist medical practitioners) have received growing attention. A systematic review by Bleich et al. [5] located 27 studies that were concerned with the effects of programs to treat people with multi-morbid CD. The studies included any that were concerned with the efficiency of health care use and spending, or patient satisfaction. The authors found that many studies reported no significant improvement on any of the aims of reducing spending, improving clinical outcomes, or increasing consumer satisfaction. They note that this is "…an especially surprising result given the expected publication bias of reporting only favourable results" [5, p. 197]. Furthermore, while the extant clinical literature has frequently emphasised the importance of targeting patients with advanced disease—to prevent "inappropriate" use of Emergency Department (ED) or other service use types—whether this is, indeed, an optimal strategy is now being questioned (e.g., [6]).

From an economic viewpoint there are two main concerns with the extant literature. First, most of the studies seek to test the effectiveness of CD Management Programs (CDMPs) only on indicators of health inputs use or health expenditures, and less often, on adherence or patient satisfaction indicators. However, a rigorous and more satisfactory cost-effectiveness analysis of CDMPs should consider health outcomes both in terms of progression of disease severity and overall survival. Second, much of the existing literature does not adequately deal with the possibility of selection bias. David, Smith-Allen and Ukert [7] note, for instance, that although several recent studies ([8]-[10]) do produce evidence of interventions "working" (e.g., reducing ER visits and hospitalisations), they do not address self-selection problems. Thus, individuals who were more likely to benefit from the intervention due to observable characteristics were also more likely to receive the intervention in these studies. Feasible approaches to address this problem—such as those provided in [11] and [12]—were not applied in these instances though.

Several recent contributions (*viz.* [6], [7] and [13]-[16]) have, however, used convincing identification strategies to estimate the causal effects of CDMPs on adherence as well as health service utilisation or expenditures. One strand of this research focuses on supply-side factors such as the attributes of providers ([7] and [13]-[15]), while another focuses on demand-side factors such as the attributes of patients ([6]; [15]; [16]). These studies may be characterised as being concerned with whether or not the marginal health productivity (MHP) of CD interventions (i.e., the extent to which these interventions improve health, at the margin) is affected either by supplier or patient heterogeneity. Studies of these kinds have started to prise open the "black box" of CDMPs, providing new insights into not only the extent to which CDMPs of various kinds "work" (i.e., have the intended

consequences) but also which supply- and demand-side aspects of them appear to be the most influential. (For a more detailed review, see the Supplementary Materials.)

In summary, a small number of recent and well-designed studies have established some instances in which a CDMP has had a causal effect on one or more targeted outcomes. These studies have been able to shed some light on the role that various sources of supply- or demand-side heterogeneity play in respect of CDMP effectiveness. Among these studies, the "outcomes" of interest have, however, been indicators of health input consumption or of adherence, and only few have been able to estimate their effect on intermediate health outcomes such as disease severity or long-term health outcomes such as disease-specific or overall mortality.

We contribute to this literature in several ways. First, we study a group of individuals who have been diagnosed with Chronic Kidney Disease (CKD) and control for the severity of their disease and the presence of comorbidities. Second, using inverse propensity weighting (IPW) methods, we study the effect of a CDMP on a range of adherence measures including the consumption of disease-specific pharmaceuticals, laboratory tests and specialist visits. This intervention is particularly interesting because it involves targeted supply-side interventions that differ in intensity according to an important demand-side influence, *viz.* the severity of disease. Third, we address the question of whether the intervention has a causal effect on indicators of health outcomes that include survival, duration to dialysis and duration to the next hospitalisation with higher CKD stage, and for Cardiovascular Disease (CVD) complications (which are regarded as the most important complication of CKD). Most importantly, we exploit the clearly-defined disease stages for CKD to test for the specific role of the demand-side characteristics of the CDMP estimating its MHP not just on "All" Stage CKD patients, but also on "Early" Stage CKD patients, and "Late" Stage CKD patients, separately. We produce evidence that this CDMP was effective across all three groups and achieved notable improvements in both treatment adherence and health outcomes. Moreover, our results present strong evidence that the largest effects of the program tended to be for Early Stage CKD patients, rather than for Late Stage CKD. This is generally true for both the adherence measures and the health outcomes we use, with the exception of time-to-dialysis where the Late CKD patients benefit more from the intervention in delaying dialysis than do Early CKD patients. Our results are broadly consistent with the notion that the MHP of CDMPs may not be maximised by targeting Late-stage disease, but by targeting patients who are at a less advanced stage of the disease's progression.

The remainder of the paper is set out as follows. First, we provide a brief description of health care financing and service delivery in Italy and in the Emilia-Romagna Region (ER) as well as of an intervention designed to improve the continuity of care for people with CKD. We then outline our empirical identification strategy to isolate the causal effect of the intervention on adherence and survival and to test whether these differ across the levels of CKD severity. In the following sections we present the econometric results and the discussion. To conclude, we comment on the possible policy implications of our results and foreshadow a research agenda to explore additional research questions with similar datasets and other methods.

2 INSTITUTIONAL SETTING AND THE PIRP PROGRAM

The Emilia-Romagna (ER) Region, in Italy's North-East, is the sixth-largest region in the country with an estimated resident population of approximately 4.45m people [17]. During the relevant period (January 2009 - December 2015) the Regional Health System was organized in 11 Local Health Authorities (LHAs) till January 2014, when they were reduced to 8 due to the merger of 4 LHAs in the

Eastern side of the Region.[1] The Region is characterised by an ageing population: in 2016 approximately one in ten people was over 75 years of age [17]. Life expectancy for men and women in 2016 was 81.2 and 85.3 years, respectively [17]. As such, the chronic diseases associated with an ageing population also represent an important challenge to the Region. Hospital admission rates for "ambulatory care-sensitive" chronic conditions (*viz.*, chronic heart failure, chronic obstructive pulmonary disease, diabetes and chronic kidney disease) exceed the Italian national standards set out in the National Outcome Evaluation Program framework [18].

The PIRP (*Prevenzione della Insufficienza Renale Progressive*) program, launched in 2004, involves a diagnostic pathway, treatment and care regimen that is codified as a set of rules for the conduct of treatment of people with CKD by General Practitioners (GP) and specialists, principally nephrologists. The rules call for a close monitoring by the GPs starting at the very early stages of the disease and for a greater involvement of nephrologists as the disease advances as well as specifying the tests and frequency of tests that should be taken at each CKD stage, broadly following the Guidelines of the (US) National Kidney Foundation [19].

From an economic perspective, PIRP can be thought as a complex CD intervention that involves a combination of demand- and supply-side measures to forestall disease progression. The program involves measures that aim to inform CKD patients about their risk factors (including non-adherence to recommended treatment) from an early disease stage; to monitor them in a timely fashion; and to refer them to the care of a nephrologist once their disease severity warrants it. The program also involves a suite of supply-side interventions that ranges from professional training for GPs, to the establishment of a clinical registry to coordinate GPs and nephrologist care, as well as administrative support to monitor CKD patients and to strengthen the supply of specialist services through additional funds to increase the ambulatory hours dedicated to CKD patients in the nephrological units.

PIRP's design strongly emphasises the need to differentiate monitoring and provision of services according to the patients' characteristics (CKD stage, sex, age, comorbidities). Accordingly, it gives priority to the education, monitoring and service delivery for early-stage CKD patients with the objective of slowing down the progression of the disease. In economic terms, the program has a clear long-term orientation as the benefits for individual patients will not be significant in the short-run and requires strong vertical integration between the payer and the providers to reduce the transaction costs of coordinating several professionals competing between themselves (GPs, nephrologists) and acting in different settings (primary care, hospitals).

3 METHODS

3.1 Identification Strategy

The objective of the econometric work is to measure the causal effect of the treatment, *viz.* enrolment in the PIRP program, on indicators of (i) treatment adherence (a health inputs measure) and (ii)

[1] The merger should have had no impact on how the PIRP programme affected the use of health services and the relevant health outcomes. Indeed, even if the merger took place on January 2014, the new organizational model of the unified LHA were to be based on a formal act (atto aziendale) of its Director to be subsequently approved by the Emilia-Romagna Region. Due to a series of administrative and political problems (see Carbone et al. 2015 for a detailed analysis), the new organizational model was proposed only in mid May 2015 and was formally authorised later on by the Region. Hence, it is unlikely that the merger – without significant organizational changes in the provision of services – might have an influence on the effectiveness of the PIRP program during the relevant period (2009-2015).

survival (a health output/outcome measure). To do so, we specify endogenous treatment effect (TE) formulations to estimate the effect of the program on (i) and (ii).

Our approach follows the well-known potential-outcomes framework [20] to estimate the causal effects of PIRP enrolment on the variables of interest:

$$Y_i = \alpha + \beta D_i + X'\vartheta + \mu_i \quad (1)$$

where α is a intercept, D_i is an indicator of treatment assignment (=1 if the patient is enrolled in PIRP; =0 otherwise), X' is a vector of control variables with coefficient vector ϑ, and u_i is the error term. The general characteristics of the model are well-known and well-documented (see, e.g. [21]-[23]). So, in the interests of space, we provide a brief explanation of the main points of the empirical strategy to identify the causal effect of the treatment. Briefly, given that assignment to the treatment (PIRP) is presumed not to be random, estimates of β via standard regression analysis (e.g., Ordinary Least Squares) produce results that represent both the estimated average treatment effect on the treated (ATET) and selection bias (cf. [21]).

The self-selection problem [23] that arises due to the (presumed) non-random assignment to D_i thus gives rise to endogeneity, rendering the $\hat{\beta}$s estimated from Equation (1) biased and inconsistent. In order to identify the causal effect of the treatment we invoke the ignorable assignment [23] or conditional independence assumption (CIA) [25],[26]. In this approach, probabilistic assignment to the treatment *D* requires that, conditional on controlling the characteristics in X_i—which cannot be variables that are, themselves, affected by treatment—the treatment assignment of the ith individual is "as good as random" and, hence, is independent of potential outcomes. This is also referred to as the un-confounded assignment condition [22] or "selection on observables" [27].

The common support assumption rules out propensity scores of zero or unity which would imply that, conditioning on X_i, some units are never treated, or always treated, respectively. Operationally, in this study we must establish that PIRP enrolees can be matched on observables with non-enrolees, and restrict our estimates of the ATE to instances where the assumption of "common support" holds. Rosenbaum and Rubin [24] refer to the set of assumptions we invoke as the "strong ignorability" assumption, and we invoke that assumption here.

While the CIA or "unconfoundedness" assumption itself is untestable [23], several arguments suggest that it is unlikely to be violated in our setting. First, as reported previously, our data come from contiguous provinces in the ER Region, across which the standard (or quality) of health care is fairly homogenous (by design) and for which there are no geographical differences in health care financing or in the organization of the supply of health care services. Second, the X_i vector available in our dataset allows us to exploit a rich set of controls not affected by the treatment which include information on demographics and health status (e.g., age, sex, a Charlson Comorbidity Index, dummies for specific comorbidities of particular relevance for CKD patients).[2] Third, we control for small-area variations by including controls for the patient's residential location in the 11 LHAs within the region. Thus, we regard the CIA as likely to hold in respect of assignment to the PIRP program, so long as the common support assumption is met. This assumption is empirically testable, and we produce evidence in the empirical results that it is satisfied for all specifications we report here.

We specify and estimate endogenous treatment effects models using inverse probability weights (IPWs) to balance patients in the treatment and control groups. A binary logit regression is employed

[2] It is worth noting explicitly that, because PIRP includes a zero co-payment entitlement to PIRP enrolees for all recommended services, we do not include prices (or copayments) in the set of X_i variables.

in which the enrolment (or otherwise) in PIRP is the outcome variable ($D_i = 1, 0$) estimated as a function of observables X_i, and we obtain the predicted values $\hat{p}(X_i)$, which are used to compute the inverse probability weights for the ATEs, following the approach described in Kaiser and Schmid [28]. For all estimates, we adjust the *p*-values using Holm-Bonferroni correction for multiple hypothesis tests [29].

Finally, we note a further assumption that is necessary for our identification procedure to be valid: namely the stable unit treatment value assumption (SUTVA) [30]. In the context of this study, the SUTVA implies that the PIRP enrolment of the *i*th individual affects only individual *i'*s outcome. This assumption seems reasonable in the current context: we do not expect one individual's enrolment in PIRP to have any (e.g., general equilibrium) effects that would violate the assumption by affecting any other individual's outcome. In the following section, we discuss further measures that were taken to examine the sensitivity of our ATE measures to various inclusion/exclusion criteria.

4 EMPIRICAL ANALYSIS

This section begins with a description of the extraction procedure to build the dataset, followed by a discussion of the methodology adopted to build the comorbidity, adherence and outcome indicators, describes how the information on the CKD stage is used to analyse the differential effect of PIRP on early and advanced CKD patients, and the role of the CKD priority level *vis a vis* other diagnoses to sustain the common support hypothesis. Finally, we present descriptive statistics for all the demographic, health status, adherence and health outcomes variables. The next section presents the econometric estimates of the effect of the treatment—enrolment in the PIRP program—on measures of survival and adherence.

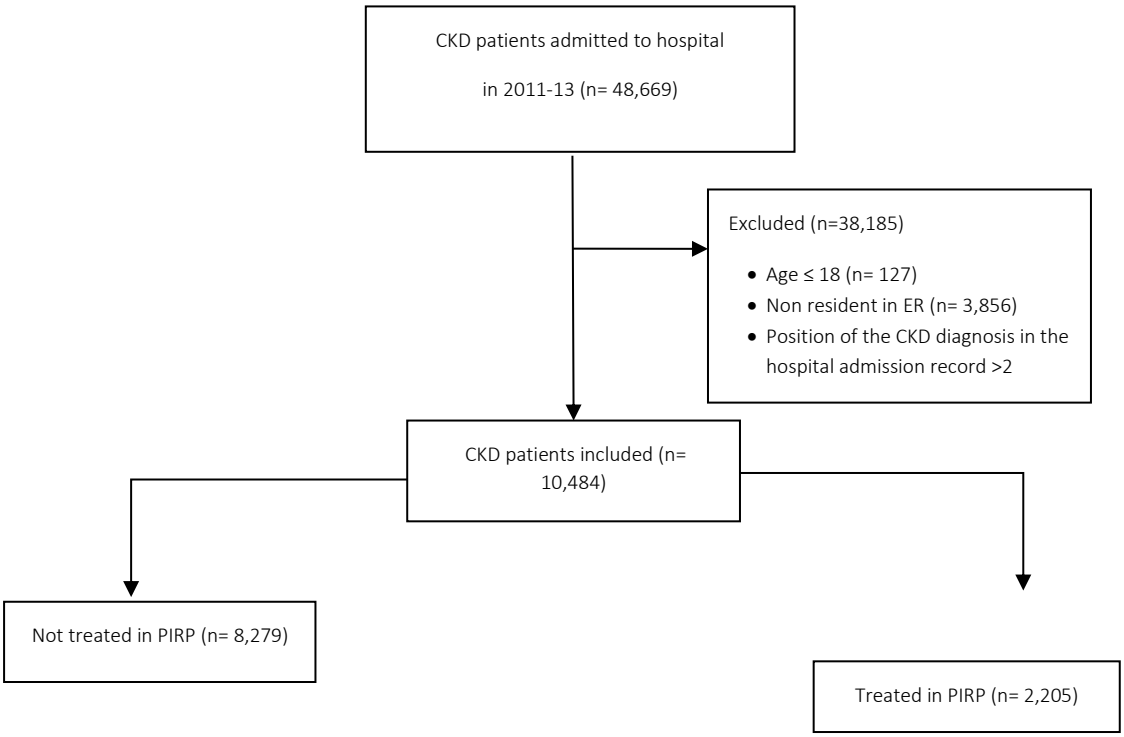
4.1 Population

Health care data were extracted for all individuals over the age of 18 years who were hospitalized in the Emilia-Romagna region between 1 January 2011 and 31 December 2013 with a primary or secondary CKD diagnosis. Specifically, ICD-9-CM codes from 585.1 to 585.6 and 585.9 were used to identify all CKD diagnoses, where the digit after the decimal place indicates the CKD Stage, or severity. CKD severity is graded in stages that run from Stage 1 to Stage 6. The code 585.9 represents CKD where the stage has not been specified. We expand on the relevance of these severity indicators in the next section of the paper.

For convenience, we refer to the first hospitalisation with primary or secondary CKD diagnosis observed in this period for the individual as their "index hospitalisation" as it is the earliest occasion where reliable information is available on the severity of the CKD progression. The inclusion criteria (see Figure 1) resulted in a sample of 10,484 individuals in the dataset.

For each individual who met the above inclusion criteria we then extracted data on a number of variables for precisely the two years preceding, and the two years following their index hospitalisations. For convenience, henceforth we will refer to these periods using the shorthand: "pre-index period" and "post-index period". The data we observe thus include observations from January 2009 and up to December 2015, with the date of the index hospitalisation determining precisely which 730 days precede and follow that event. We use the pre-index period to construct indicators of patient comorbidities and use the post-index period to compute indicators of adherence to the recommended treatment regimen, which varies by disease stage. We use data for the post-index period to compute and model the adherence indicators and outcome measures that are described in detail below.

Figure 1 – Inclusion and exclusion criteria



The variables available to us, at the patient level, include the patient's age, gender, LHA of residence, whether and when the patient is PIRP-enrolled, the number of hospital admissions, the reason(s) for admission to hospital - from which we build the Charlson Comorbidity Index (more on this below) -, the position of the CKD diagnosis in the hospital record, the CKD stage recorded at hospital admission (six stages, in total), the number of visits to specialists (disaggregated by discipline, e.g. nephrologist, cardiologist), numbers of laboratory tests taken (by type), consumption of pharmaceuticals used to treat CKD and four common co-morbidities (hypertension, hyperlipidemia, diabetes and cardiovascular diseases), an indicator of dialysis, and mortality.

In our dataset, dates were included alongside all recorded events (e.g., hospital admissions, specialist visits, first dialysis, mortality) and for all laboratory tests and pharmaceuticals consumption data. We are thus able to create indicators of consumption and adherence in respect of these measures, with reference to treatment guidelines. We provide further details below. Briefly, in respect of specialist visits and laboratory tests, we are confident of a "one-to-one" relationship between purchases and consumption, as the data are recorded by pharmacists and specialists who are only remunerated for recorded services. For pharmaceuticals, though, it is possible that not all purchased pharmaceuticals are actually consumed. For the purposes of this study, though, we assume that purchases or, more specifically the "medication possession ratios" (MPRs) we construct from them, represent a reasonable proxy for consumption [31].

These data enable us to examine guidelines adherence for at least two years prior to, and after, the index hospitalisation and to examine survival outcomes, including time to dialysis for those patients ever dialysed, and survival (time-to-death) times for all patients in the dataset. As is explained in detail below, we exploit these two-year windows before and after the "recruitment" window, to create

indicators of comorbidities and of (relatively) long-term adherence (i.e., over a 2-year period) as well as to create time-to-event measures of outcomes after the index hospitalization.

4.2 Chronic Kidney Disease (CKD) Priorities and Stages

The National Kidney Foundation Guidelines [19] provide a classification of CKD and recommendations how its management should change with disease severity. The Guidelines, which are endorsed by the PIRP program, separate CKD into six clearly-defined disease stages: Stage 1 is the classification given to lowest-severity of disease, and the successive stages, which go through to Stage 6 (i.e., end-stage renal disease), represent increasing disease severity. Table 1 shows how the clinical management of CKD is recommended to adjust to different levels of disease severity.

Table 1 Operational Guidelines for the Management of Chronic Kidney Disease (CKD)

	Minimum Recommended Annual Consumption, by service and CKD Stage			
	1-2	**3**	**4**	**5-6**
Nephrologist visits	1	2	4	8
Laboratory tests				
Urinalysis	1	2	4	8
Creatinine analysis	1	2	4	8
Microalbuminuria analysis	1	2	4	8
Proteinuria analysis	1	2	2	4
Phosphorous analysis	-	1	2	4
Calcium analysis	-	1	2	4
Parathyroid analysis	-	1	2	4

The facts that CKD has clearly-defined stages, and that the recommended treatment changes with those stages, are important to our empirical strategy for several reasons. First, the PIRP program encourages the enrolment of CKD patients when their disease severity warrants it: specifically, at CKD Stages 3-4. Patients whose disease severity meets the CKD Stages 1-2 definitions are considered too early in their disease to warrant enrolment in the PIRP program; it is recommended that their care be coordinated by GPs. Conversely, patients with CKD at Stages 5-6 have severe disease, which is not the target of the PIRP program; and they should be managed intensively by nephrologists but not referred into PIRP. Second, notwithstanding the targeted nature of the PIRP program, people are still enrolled in it at all disease stages. This enables us to consider whether and to what effect, enrolment in the program has on adherence and outcomes at different levels of disease severity, i.e. to determine whether its average treatment effects (ATEs) vary by disease stage. This includes addressing the interesting, and policy-relevant, question of whether or not those ATEs are larger or smaller at different CKD stages.

The hospital records available to us contain up to 15 Major Diagnostic Categories (MDCs): these are typically listed in order of the hospital clinicians' opinion about the relative importance of each diagnosis to the reason for this hospital episode. We decided to limit our inclusion, in this work, only to those patients with a relatively high-priority diagnosis of CKD (i.e., first or second diagnosis in this

paper; and first through fifth position in sensitivity analyses). This decision was taken following discussions with specialist nephrologists, that outlined that CKD patients are very often multimorbid patients with also cardio-vascular and endocrinological conditions. Hence, the selection of patients with CKD as a low-level diagnosis would have broadened up the sample to include patients for whom the role of the nephrologists and of the PIRP programme were likely to play a marginal role in influencing the main outcomes of interest. Moreover, we were advised that each patient's CKD stage was likely to be accurately reported only when CKD was considered a high-priority in the treatment of the patient, as indicated by its place in the hospital record. In sensitivity analyses, we included CKD patients with hospital record priorities from one to five: the results are qualitatively similar so, although they involve larger numbers of individuals, we do not report them here, but they are available upon request.

Subject to the foregoing inclusion criteria, we then estimated the ATEs for the PIRP program on all patients who met the inclusion criteria. We refer to this group as "All CKD" and estimates of the ATEs therefore represent ATEs for CKD patients at any stage of disease. It is worth emphasising that this categorisation includes patients for whom the CKD stage was coded as "9": this is not, in fact, a CKD stage but reflects uncertainty on the part of the diagnosing practitioner (at the time the diagnosis is recorded on the hospital record) as to the specific CKD stage. Upon investigation, we established that this coding most frequently was confirmed in the PIRP register as being CKD stages 3 or 4. We included all such patients in the first "All CKD Stages" category. This also explains why the number of observations in this category is always super-numerate with respect to the sum of the next two subsets of patients for whom we estimate ATEs.

To examine the effect of the PIRP intervention by CKD stage, we partition the latter dataset to estimate ATEs for PIRP for only individuals with "Early CKD" (which we define as CKD stages 1-3) and with "Late CKD" (which we define as CKD stages 4-6), dropping any individual with a CKD stage coded as nine (see above). To estimate the ATEs for Early CKD, we excluded three LHAs (Parma, Modena and Forli), because their inclusion frequently led to a violation of the common support assumption.

Finally, it is worth noting that although CKD Stages are clearly defined by the relevant Guidelines [19] diagnoses of stages are, of course, still prone to measurement error. The errors arise due to misclassifications of the latent stage by treating doctors: and we know, from existing empirical evidence, that the frequency of miscoding tends to decline in the severity of the disease [32]. The critical question in this study is whether any systematic differences in measurement error—with respect to CKD stage—attends PIRP-enrolled and PIRP non-enrolled patients. If there were systematic differences in the CKD stages assigned to PIRP and non-PIRP patients, this would be of particular concern. The resulting ATE estimates would also then be subject to bias. There is no basis, we argue, to suspect systematic differences in measurement error between the treatment and control groups in this study. The CKD stages for all patients in the sample are supplied by the same hospital staff, irrespective of whether or not they are PIRP participants. Furthermore, PIRP participant status is not something that is routinely revealed to treating hospital practitioners either; so our expectation is that measurement error is orthogonal with respect to PIRP participation.

4.3 Adherence Indicators

We use data on the purchase of specialist visits, laboratory tests and pharmaceuticals, along with dates of consumption, to construct adherence measures for each service type. To do this we take the individual's consumption of treatment, D_j, over time period *T*, as a proportion of the regimen-recommended quantity R_i:

$$0 < \frac{\sum_{t=1}^{T} D_{ij}}{R_T} \leq 1 \qquad (2)$$

where T is the number of days in the time period over which adherence is defined and j indexes the specific treatment type (e.g., creatinine testing). In words, we create a ratio that takes values between zero and unity, defined over the guideline-specific time period, T, for each service type for which adherence guidelines are available[3]. Medication adherence indicators of this kind are sometimes referred to as Truncated Medication Possession Ratios (TMPRs) and have a long history in the drug adherence literature. See, e.g., [33] and [34]. The richness of our dataset also enables us to compute adherence indicators (AIs) in the form of TMPRs for each class of drug.

Along with the implicit assumption that pharmaceutical purchases present a suitable proxy for consumption, two further assumptions apply to these measures *viz.*: it is assumed that (i) prescriptions are not made in diagnostic error (i.e., that the disease for which they are prescribed is not due to a false-positive diagnosis), and (ii) the diseases for which these compounds are prescribed are chronic and hence, do not self-resolve and are not "cured" by treatment. While either assumption may be violated for some individuals, we do not expect this to be a large source of measurement error.

To compute adherence indicators (AIs) for specialist nephrologist visits and laboratory tests, we apply an operational version of the guidelines issued in 2009 by a group of nephrologists, GPs and epidemiologists appointed by the largest LHA of the ER Region and endorsed by the nephrological units in charge of the PIRP program (see [35]) to make similar computations to those described at Equation (2). We compare the actual numbers of specialist visits and laboratory tests consumed to the consumption levels that are indicated by the guidelines at each stage of CKD. We express this as a ratio that, as with the MPR, has a lower bound of zero and an upper bound of unity. Specifically, patients who consume at, or above, the recommended frequencies of these services are deemed to be "fully compliant" with the guidelines and are designated an AI value of 1, for that service type. For instance, when the recommended minimum number of nephrologist visits for the individual's (known) CKD stage is two per annum, T=365: an individual who saw a nephrologist twice or more would is assigned an adherence indicator = 1; and an individual with the same CKD stage who visited the nephrologist only once in that period is assigned adherence indicator =0.5, and so on.

These guidelines are the basis of the denominators, R_j, we employ to compute AIs for each of the *j* adherence measures. More specifically, they provide the minimum quantities of specialist nephrologist visits and laboratory tests (by type) for each CKD stage from one to six (where higher stages represent greater disease severity), as shown in Table 1.

4.4 Comorbidity indicators

Comorbidity was assessed as an unweighted count of the 17 conditions of the Charlson Comorbidity Index ([36], [37]) in the pre-index period. In our dataset it takes on values between 0 and 10, with zero indicating no co-morbidities and the Index increasing with an increasing number of conditions. We use this Index as one of the control variables in our application of IPW.

[3] A continuous indicator of adherence instead of a dichotomous variable across a given threshold allows to estimate the marginal effect of PIRP inclusion on adherence over the whole distribution of values of adherence, even far from the standard threshold. Such effect, even if it does not lead to a level of compliance that is regarded as appropriate, might bring about significant changes in the health outcomes of chronic patients.

Second, we examine the individual's use of any drug that is prescribed exclusively for the treatment of hypertension, hyperlipidemia, diabetes or cardiovascular disease in the two years before index hospitalization and use this to characterize each patient as being affected by one or more of the main comorbidities associated with CKD. For all individuals classified as having one of the CDs, we compute the AIs described in the preceding section for the medications associated with its treatment.

4.5 Adherence and Health Outcome Indicators

Our primary adherence measures—which conceptually are health production inputs (see, e.g. [38],[39])—are the (specialist visit and laboratory tests) service and pharmaceutical adherence indicators defined in the previous section. Our primary outcome measures are (i) mortality (time-to-death), (ii) CKD disease progression (time-to-higher stage; time-to-dialysis) and (iii) time-to-hospitalisation for cardiovascular (CVD) problems. The null hypotheses are that the treatment—PIRP enrolment—has no effect on adherence or survival. Conceptually, the PIRP program is designed to improve health status by enhancing adherence. In economic terms, we conceive of the program's potential effect on adherence as being to influence patients to engage more efficiently in health production by following recommendations that increase, up to the recommended threshold, the service and time inputs consumers use in their health production functions.

We operationalise time-to-event (survival) models by estimating endogenous treatment time-to-next hospital-admission with higher CKD stage, time-to-next-hospitalization for CVD problems, time-to-dialysis and time-to-death models. In each instantiation we specify a Weibull censoring function, logit treatment assignment model and a weighted mean outcomes model, where the estimated IPWs derived from the logit estimates provide the weights. The econometric models for adherence are also operationalised as weighted mean outcome models, where the means are weighted by the IPWs estimated on an endogenous logit treatment model of treatment assignment.

4.6 Common Support and Inclusion Criteria

As unbiased treatment effects are only estimable when the common support assumption is satisfied, it is important to examine the normalised differences between the treatment and control groups. The common support assumption was tested using a formal test of it, *viz.* the over-identification test developed by Imai and Ratkovic [40]. In this test the null hypothesis that the covariates are balanced with a test statistic that is distributed as χ^2. In the Tables of results, we report the level of significance at which the null hypothesis of no common support may be rejected by the result. We comment specifically on whether the null hypothesis of common support is rejected in our presentation of the results.

Briefly, though, our results on common support show that for "All CKD Stages" and "Late CKD" the null of common support cannot be rejected, at the one per cent level, for any of the results reported in the next section. For "Early CKD", the results on common support are, however, mixed: while the common support assumption for this subgroup is always passed at the 10 per cent level, it is not passed at higher levels of significance in a large minority of cases. In such cases, we still report the results on the "Early CKD" group, but interpret the results with caution. Nevertheless, in such instances, comparisons of the "All CKD stages" results with those of the "Late CKD" group is informative. Since the common support assumption for these two groups is never rejected at the one per cent level, a comparison between the two provides a further way of comparing the MHP of the intervention (via the ATEs) of people with Late-stage against those of patients at all disease stages. Essentially, doing so always results in a similar logical conclusion to that which would be drawn by considering the "Early CKD" results on their own.

4.7 Descriptive statistics

Table 2 shows the descriptive statistics for all the demographic, health status, adherence and health outcomes variables. The CKD patients enrolled in PIRP – especially those at an early stage - are on average younger, more likely to be male (as the progression of CKD at all stages is typically much faster for males), with a lower Charlson Comorbidity Index and number of hospital admissions in general, but more likely to have CKD related comorbidities (hypertension, hyperlipidemia, diabetes and CVD) and hospitalizations.

As for the indicators of adherence to the guidelines, the PIRP patients—again especially for the early CKD stages—have higher levels of compliance for drug consumption related to the main comorbidities, laboratory tests and nephrologist visits. Analogous differences can be observed for the main health outcomes and especially for the mortality rates and for the time to dialysis.

In next section we will show that these significant differences in adherence and in health outcomes can be attributed only partly to difference in the patients' characteristics and/or to a self-selection effect and demonstrate a causal effect of PIRP on both adherence and health outcomes.

In respect of the time-to-event and mortality data towards the bottom of Table 2, some points are noteworthy. First, the time-to-event data are based on the whole sample, as opposed to conditioning on the event having actually occurred. For example, time-to-death is calculated over those who died within 730 days as well as those who did not. We computed (but do not report for reasons of space) descriptive statistics on the conditional means of these time-to-event data and found that, with two exceptions, they typically were quite close in magnitude to the unconditional means reported in Table 2. The two exceptions were for the two least common events, *viz.* dialysis and death. On these indicators the durations fell to more than half the unconditional values reported in Table 2. Second, while deaths are not as common as other events we observe (such as cardiovascular hospital admissions or changing CKD stage), the mortality rates that we observe are nevertheless very high in this population: by the end of follow-up at two years approximately 37% of non-PIRP patients and 20% of PIRP patients in our dataset had died. These two points together illustrate that the sentinel events we observe over a two-year period post-hospitalisation, represent fairly common events in this patient population, which has an average age of over 75, has been diagnosed with CKD, and often carries substantial comorbidities.

5 CAUSAL EFFECTS OF PIRP

5.1 Adherence

We hypothesised that the mechanism via which the PIRP program may change indicators of health status is via its effects on AIs for the recommended treatments. In this section we report estimates of the causal effects of PIRP on specialist visit adherence, a range of pharmaceutical adherence indicators, and laboratory testing using the AIs described in Section 4. For the latter, we report adherence by test type as well as estimating an aggregate adherence measure for adherence to the testing regimen recommended for CKD patients, by disease stage. To enable comparisons with the treatment effect results that will be reported for health outcomes, we use the same inclusion criteria in these models. Recall that, rather than imposing an adherence threshold (e.g., specifying a proportion that represents good/poor adherence), we test for adherence AI measures as dependent variables and testing the null hypothesis of a zero effect of the PIRP program on them. Recall that, for all services—pharmaceuticals, laboratory tests and specialist nephrologist visits—all AIs are specifically defined with respect to the CKD stage, following the PIRP guidelines.

Table 2 - Descriptive statistics

Variables	Total (10,484 pts)*				Early CKD stage (5,471 pts)*				Late CKD stage (3,392 pts)*			
	n-Pirp (8,279) mean	Pirp (2,205) mean	diff/t-test diff	t-test	n-Pirp (4,615) mean	Pirp (856) mean	diff/t-test diff	t-test	n-Pirp (2,325) mean	Pirp (1,067) mean	diff/t-test diff	t-test
%Sex (female=1)	0.45	0.35	0.10	8.49	0.44	0.32	0.12	6.81	0.43	0.37	0.06	3.3
Age	75.53	72.1	3.43	9.51	76.32	70.62	5.70	10.01	71.89	72.31	-0.42	0.75
%Age>70	0.74	0.69	0.05	4.92	0.76	0.65	0.11	6.66	0.63	0.69	-0.05	2.82
Charlson index	1.35	1.15	0.20	5.86	1.33	1.19	0.14	2.66	1.32	1.10	0.22	4.13
% Previous hypertension	0.79	0.87	-0.08	8.88	0.78	0.84	-0.07	4.33	0.82	0.9	-0.07	5.55
% Previous hyperlipidemia	0.37	0.54	-0.16	14.07	0.37	0.54	-0.17	9.13	0.40	0.55	-0.15	8.19
% Previous diabetes	0.23	0.31	-0.08	7.77	0.23	0.28	-0.05	3.19	0.24	0.35	-0.11	6.48
% Previous CVD	0.48	0.56	-0.08	6.42	0.47	0.48	-0.01	0.71	0.52	0.64	-0.11	6.2
# Previous hospitalization	1.48	1.16	0.32	6.85	1.43	1.19	0.24	3.42	1.46	1.07	0.39	5.49
# Previous hosp for CKD	0.27	0.32	-0.05	2.42	0.19	0.27	-0.08	2.93	0.47	0.40	0.07	1.87
# Previous nephro visit	1.60	3.56	-1.96	19.1	0.93	2.18	-1.25	12	3.45	5.02	-1.58	7.02
% Population resid LHA1	6.42	4.76	1.66	2.91	7.11	6.07	1.03	1.09	6.62	4.68	1.94	2.21
% Population resid LHA2	8.99	3.76	5.23	8.12	7.37	1.75	5.62	6.15	11.66	4.97	6.69	6.19
% Population resid LHA3	8.30	12.11	-3.81	5.54	8.36	11.21	-2.85	2.7	8.77	13.5	-4.72	4.22
% Population resid LHA4	14.91	9.02	5.88	7.16	17.16	6.89	10.27	7.65	13.76	12.00	1.77	1.41
% Population resid LHA5	21.09	20.5	0.59	0.61	23.92	23.01	0.91	0.57	22.37	22.02	0.34	0.22
% Population resid LHA6	1.64	2.68	-1.03	3.19	1.30	2.57	-1.27	2.81	2.62	2.44	0.18	0.32
% Population resid LHA7	9.95	9.84	0.11	0.16	5.68	5.61	0.07	0.08	9.20	7.59	1.61	1.55
% Population resid LHA8	10.83	6.98	3.85	5.36	8.49	7.36	1.13	1.10	9.08	4.31	4.77	4.88
% Population resid LHA9	5.37	5.53	-0.16	0.29	7.37	7.59	-0.22	0.23	3.87	4.78	-0.91	1.23
% Population resid LHA10	4.72	14.20	-9.48	15.98	6.28	17.06	-10.77	10.8	1.68	13.03	-11.35	14.16
% Population resid LHA11	7.75	10.61	-2.86	4.31	6.96	10.86	-3.91	3.98	10.37	10.68	-0.32	0.26
Adher nephro visits	0.32	0.67	-0.35	33.04	0.26	0.7	-0.44	29.72	0.44	0.64	-0.21	13.54
Adher lab tests	0.35	0.50	-0.14	20.51	0.37	0.56	-0.19	19.1	0.32	0.44	-0.12	12.38
Adher med hypertension**	0.85	0.92	-0.07	11.27	0.84	0.91	-0.07	6.42	0.87	0.94	-0.06	6.89
Adher med hyperlipidemia**	0.64	0.68	-0.04	3.94	0.65	0.73	-0.09	5.56	0.62	0.64	-0.02	1.02
Adher med diabetes**	0.68	0.72	-0.03	2.09	0.70	0.74	-0.04	1.9	0.65	0.69	-0.04	1.87
Adher med CVD**	0.74	0.78	-0.04	4.7	0.70	0.72	-0.02	1.07	0.81	0.83	-0.02	1.46

Table 2 Cont'd

Variables	Total (10,484 pts)*				Early Ckd Stage (5,471 pts)*				Late Ckd Stage (3,392 pts)*			
	n-Pirp (8,279) mean	Pirp (2,205) mean	diff/t-test diff	t-test	n-Pirp (4,615) mean	Pirp (856) mean	diff/t-test diff	t-test	n-Pirp (2,325) mean	Pirp (1,067) mean	diff/t-test diff	t-test
time to higher CKD (days)	478.01	501.25	-23.24	3.33	513.36	576.12	-62.76	6.15	381.06	411.3	-30.24	2.68
time to next CVD hospit (days)	228.99	271.81	-42.81	4.79	241.17	281.84	-40.68	2.52	213.18	263.89	-50.72	4.22
time to dialysis (days)	511.36	551.09	-39.73	5.92	556.2	653.81	-97.62	10.44	414.76	451.95	-37.19	3.32
time to death (days)	554.2	650.41	-96.22	16.09	560.31	665.61	-105.3	11.38	453.54	640.77	-87.23	9.65
mortality 30 days	0.05	0.01	0.04	7.56	0.05	0.01	0.04	4.91	0.05	0.01	0.04	5.19
mortality 180 days	0.18	0.06	0.12	13.62	0.17	0.04	0.13	9.98	0.18	0.07	0.10	7.92
mortality 365 days	0.26	0.11	0.14	14.56	0.25	0.09	0.16	10.19	0.26	0.13	0.13	8.73
mortality 730 days	0.37	0.20	0.16	14.65	0.36	0.16	0.19	11.06	0.37	0.23	0.14	8.44

*The sum of the patients with Early CKD stage (5,471) and of those with Late Ckd Stage (3,392) does not add up to the total of 10,484 patients because there are 1,621 patients whose CKD stage at the index hospital admission has been recorded as "indeterminate". Based on the data of the PIRP Clinical Register, the modes of the stage for the patients with code "indeterminate" are Stages 3 and 4.

** The number of observations for the adherence indicators is lower (see the percentages in Table 2A) because they have been calculated only for the patients whose comorbidities was detected in the two years before the index hospitalizations.

Table 3: Estimated Causal Effects of the PIRP Program on Indicators of Specialist Visit, Pharmaceuticals, and Pathological Testing Adherence

	All CKD stages				Early CKD				Late CKD			
Adherence Indicators for:	ATEs (S.E.s)	PO Means (S.E.s)	ATEs/PO Means (%)	*n*, overid fail?	ATEs (S.E.s)	PO Means (S.E.s)	ATE/PO Means (%)	*n*, overid fail?	ATEs (S.E.s)	PO Means (S.E.s)	ATE/PO Means (%)	*n*, overid fail?
Specialist Visits												
Nephrologist visits	0.35*** (0.009)	0.32*** (0.01)	109%***	8,863[ttt]	0.44*** (0.02)	0.24*** (0.007)	183%***	3,860 [t]	0.20*** (0.01)	0.44*** (0.009)	45%***	3,392[ttt]
Pharmaceuticals												
Hypertension drugs	0.05*** (0.007)	0.86*** (0.003)	6%***	9,232[ttt]	0.04*** (0.01)	0.85*** (0.005)	4%***	3,324 [tt]	0.04** (0.01)	0.88*** (0.006)	5%***	3,142[ttt]
Hyperlipidemia drugs	0.02** (0.01)	0.64*** (0.006)	3%**	4,471[ttt]	0.07*** (0.02)	0.66*** (0.009)	11%***	1,628[ttt]	0.01 (0.02)	0.63*** (0.01)	0%	1,589[ttt]
Diabetes drugs	-0.02 (0.02)	0.69*** (0.008)	0%	2,647[ttt]	-0.003 (0.03)	0.71*** (-0.01)	0%	940[ttt]	0.01 (0.02)	0.66***	0%	945[ttt]
Cardiovascular disease drugs	0.03*** (0.01)	0.74*** (0.004)	4%***	6,748[ttt]	0.006 (0.03)	0.71*** (0.007)	0%	2,243[ttt]	-0.0005 (0.01)	0.82*** (0.007)	0%	2,504[ttt]
Laboratory Tests												
Creatinine tests	0.14*** (0.009)	0.67*** (0.005)	21%***	8,863[ttt]	0.20*** (0.01)	0.70*** (0.007)	29%***	3,860[ttt]	0.13*** (0.01)	0.59*** (0.009)	22%***	3,392[ttt]
Urine tests	0.12*** (0.01)	0.42*** (0.005)	29%***	8,863[ttt]	0.24*** (0.02)	0.47*** (0.008)	51%***	3,860[ttt]	0.11*** (0.01)	0.27*** (0.008)	41%***	3,392[ttt]
Microalbuminuria tests	0.09*** (0.009)	0.13*** (0.004)	69%***	7,757[ttt]	0.18*** (0.02)	0.14*** (0.005)	129%***	3,860 [t]	0.07*** (0.01)	0.07*** (0.005)	100%***	2,286 [tt]

Table 3 (Cont'd)

	All CKD Stages				Early CKD				Late CKD			
Adherence Indicators for:	ATEs (S.E.s)	PO Means (S.E.s)	ATEs/PO Means (%)	*n*, overid fail?	ATEs (S.E.s)	PO Means (S.E.s)	ATE/PO Means (%)	*n*, overid fail?	ATEs (S.E.s)	PO Means (S.E.s)	ATE/PO Means (%)	*n*, overid fail?
Protein tests	0.18*** (0.01)	0.37*** (0.005)	34%***	8,863ŧŧŧ	0.27*** (0.02)	0.38*** (0.008)	71%***	3,860ŧŧŧ	0.12*** (0.01)	0.35*** (0.008)	34%***	3,392ŧŧŧ
Calcium tests	0.28*** (0.01)	0.50*** (0.007)	56%***	6,539ŧŧŧ	0.38*** (0.02)	0.45*** (0.01)	84%***	2,182ŧŧŧ	0.18*** (0.02)	0.57*** (0.01)	32%***	3,392ŧŧŧ
Phosphorous tests	0.02*** (0.005)	0.03*** (0.002)	67%***	6,539ŧŧŧ	0.01** (0.008)	0.02*** (0.003)	50%**	2,182ŧŧŧ	0.01 (0.008)	0.05*** (-0.004)	0%	3,392ŧŧŧ
Parathyroid tests	0.29** (0.01)	0.27*** (0.006)	107%***	6,539ŧŧŧ	0.35*** (0.02)	0.19*** (0.009)	184%***	2,182ŧŧŧ	0.20** (0.02)	0.36* (0.009)	56%***	3,392ŧŧŧ
Aggregate Measure (All laboratory test adherence)	0.15*** (0.007)	0.35*** (0.003)	43%***	8,863ŧŧŧ	0.21*** (0.01)	0.36*** (0.005)	58%***	3,860 ŧ	0.12*** (0.009)	0.32*** (0.005)	38%***	3,392ŧŧŧ

Notes: (i) The columns referring to All CKD stages include all patients with a CKD diagnosis whose hospital record placed that diagnosis in the first or second position (of up to 15 Major Diagnostic Categories (MDCs)); (ii) the columns referring to Early CKD include all patients with CKD Stages one to three whose hospital record placed that diagnosis in the first or second position of MDCs used, excluding the LHAs of Forli, Modena and Parma due to violations of the common support assumption ; (iii) the columns referring to Late CKD include all patients with CKD Stages four to six whose hospital record placed that diagnosis in the first or second position of MDCs used; (iv) one, two and three asterisks ("*") indicate statistical significance at the ten, five and one per cent levels, respectively (v) ATEs are the estimated Average Treatment Effects; (vi) S.E.s are the standard errors; ATE/PO Means (%) are computed by dividing ATEs by PO means and multiplying by one hundred and are afforded the same level of statistical significance as the ATE estimates; (vi) *n* is the number of observations; (vi) overid reports the results of Imai and Ratkovic's (2014) overidentification test, where one, two and three daggers ("ŧ") indicate that the common support hypothesis cannot be rejected at the ten, five and one per cent levels respectively.

Table 3 presents the main results on AIs. For ease of interpretation, we also present the ATEs as proportions (%) of the Potential Outcome Means (POM). Moreover, we report the rejection of the null hypothesis of no common support using daggers ("ŧ"), in superscript, that are presented alongside the sample size (*n*) for each set of results: one, two and three daggers indicate that the hypothesis may be rejected at the ten, five and one per cent levels of significance, respectively. Note that the common support assumption generally cannot be rejected even at the one per cent level. The exceptions are for Early Stage CKD patients in respect of nephrologist visits, hypertension drugs and the aggregate testing measure. Thus, some caution is warranted in respect of the ATEs on these measures.

Some of largest overall effects of the PIRP intervention on adherence are on specialist nephrologist visits adherence: and the estimated ATEs on this measure are very large. The estimated ATE for All CKD Stages is that the intervention increases compliance by 0.35 for this group, which represents a 109% increase from baseline (i.e., PO Means) adherence of 0.32, for a total adherence rate of 0.67. For the Late CKD group, the estimated increase in compliance—from a higher baseline rate of0.44—is 0.20, which represents a 45% increase in compliance due to PIRP. The largest estimated ATE for PIRP is for the Early CKD group, where PIRP is estimated to increase adherence from a very low baseline adherence estimate of 0.24 to 0.68 (an ATE of 0.44), which represents an estimated increase of 183%. Note too that, although the common support assumption is rejected below the 10 per cent level, a comparison of the All CKD and Late CKD results supports the notion that the largest adherence effect is to be found in the Early CKD group.

The next section of Table 3 reports the ATEs on adherence to pharmaceutical dosing guidelines. These produce mixed results: for hypertension and hyperlipidemia drugs, the effects of the intervention on adherence are relatively small, but statistically significant (the exception is hyperlipidemia drugs in the Late CKD group). The hyperlipidemia adherence results suggest that the entirety of the gain in adherence comes from an 11% increase in adherence for the Early CKD group, with the zero estimated increase for the Late CKD group diluting the effect for the All CKD Stages group to an approximately 3% increase. For cardiovascular disease drugs, a statistically significant increase of 4% is estimated for All CKD Stages, but the estimates for Early and Late CKD patients are statistically insignificant.

The final panel of Table 3 reports the results on laboratory test AIs. The final row of the Table reports the results on the aggregate laboratory test AI, which is a measure of adherence across all other test types. The estimates suggest a large improvement in adherence for All CKD Stages, of 38%, and the effect for Late CKD is, as a proportion, the same (although the ATE for the latter is 0.12 while the ATE for the former is 0.14). The largest percentage increase in adherence is once again for the Early CKD group, for whom the effect of the intervention is estimated to be approximately 58%: the baseline (i.e., POM) compliance rate of 0.36 is increased by 0.21 points overall. Turning to the results of adherence on individual tests, the main result to note is that the largest results of the intervention, in percentage terms, are universally for the Early CKD group. The estimated adherence effects are very large: all but one of them (creatinine tests) represents an intervention effect of 50% or more improvement in adherence. For parathyroid testing, the intervention results in a very large improvement for this group of approximately 184%: from a low baseline of 0.19, adherence with the PIRP intervention increases by an estimated 0.35 points.

Together, the results in Table 3 suggest four important things about the effect of the PIRP program on adherence. First, there is strong evidence that larger absolute and proportional gains accrue to the intervention in Early CKD rather than Late CKD. Second, an important outcome of the PIRP intervention is a very large increase in adherence to the guidelines in respect of nephrologist visits. Third, laboratory testing adherence increases quite dramatically in tandem with the increased adherence with the recommendations on specialist visits suggesting a high degree of complementarity

between nephrologist visits and disease monitoring. Fourth, the program does increase adherence with the recommended pharmaceuticals regimens, but the results are relatively small, and do not apply to all drugs. In the specific case of diabetes medications, one may suggest two, related, explanations. Note that the POM for diabetes medications suggest fairly good extant rates of compliance without the intervention (of between 0.66 and 0.71); furthermore, diabetic patients are also subject to disease-specific guidelines for their treatment together with a well-developed CDMP. For these two reasons the marginal productivity of PIRP in respect of adherence may be negligible: or, as we estimate, zero.

5.2 Time-to-Event and Probability of Death

Table 4 reports the results of estimating the causal effect of the PIRP program on the health outcome indicators we constructed from the dataset. These include deaths within two years, within one year and within 180 days as well as progression to dialysis, a more severe stage of CKD, and hospitalisations for CVD. The results at the top of Table 4 report the estimated ATEs for PIRP on the probability of deaths within these time-periods. The remainder of the Table reports the results of time-to-event (survival) models for deaths within these three timeframes, and for dialysis, CKD hospitalization with higher CKD stage and CVD hospitalisation.

Notice first that the common support assumption holds at the one per cent level for all of the estimates in Table 4 on “All CKD stages” and “Late CKD”. For “Early CKD”, though, the first six rows of the Table show that the common support assumption holds only at the ten per cent level (in each instance, the *p* value was 0.08). Second, most of the ATEs reported in Table 5 are statistically significant and all have the expected signs.

The effect of the PIRP program on the probability of dying is substantial: for the estimates presented for All CKD stages, the probability of dying within two years is 39% lower for PIRP participants than non-PIRP participants and at one year and 180 days the estimates suggest a 48% and 59% lower probability of death. Interestingly, the reduction in the chance of dying within two years for PIRP participants is estimated to be reduced by 46% for people who have Early CKD and by 35% for people with Late CKD. The estimated ATEs for death within one year or 180 days are, as proportions, largest for those PIRP participants with Late CKD: the likelihood of these patients dying is reduced by 44% and 59%, respectively, compared with those patients who were not enrolled in the PIRP program.

The survival (time-to-death) results in Table 4 are generally consistent with the death probability results: PIRP participants have better survival than non-participants at two years and one year, while the evidence of an effect at 180 days is somewhat mixed, with only a few results being statistically significant at the 10 per cent level. Focusing on one- and two-year survival, the results suggest substantial survival benefits for PIRP participants. When one does not distinguish between Early and Late CKD, PIRP participants experience mean survival gains of 34% and 21% (amounting to approximately 48 and 62 days) respectively. Distinguishing by CKD stage, however, reveals substantial differences in the survival effects of PIRP. For Early CKD, PIRP participants experience an average 41 per cent increase in survival time (equivalent to approximately 105 days) while PIRP participants with Late CKD experience a relative gain of only 14% (approximately 42 days). Of course, shorter survival durations are to be expected for CKD patients with more Late disease, irrespective of the intensity of treatment or levels of adherence. Nevertheless, these results on survival are consistent with the superior adherence estimates produced on Early CKD patients in the preceding section.

The time-to-dialysis results in Table 4 suggest that PIRP enrolment also delays progression to dialysis. Across all CKD stages, the estimated effect is approximately 26% of the POM and the estimate for Early

Table 4: Estimated Causal Effects of the PIRP Program on Health Outcome Indicators (Probability of Dying, Time-to-Death, Time-to-More Late CKD stage and Time-to-Hospitalisation for Cardiovascular Disease)

	All CKD stages				Early CKD				Late CKD			
Survival Indicator:	ATEs (S.E.s)	PO Means (S.E.s)	ATE/PO Means (%)	*n*, overid pass?	ATEs (S.E.s)	PO Means (S.E.s)	ATE/PO Means (%)	*n*, overid pass?	ATEs (S.E.s)	PO Means (S.E.s)	ATE/PO Means (%)	*n*, overid pass?
Probability of Dying												
Within 2 years	-0.14*** (0.01)	0.36*** (0.003)	-39%***	10,484[ttt]	-0.16*** (0.02)	0.35*** (0.008)	-46%***	3,860 [t]	-0.13%*** (0.02)	0.37*** (0.01)	-35%***	3,392[ttt]
Within 1 year	-0.12*** (0.008)	0.25*** (0.003)	-48%***	10,484[ttt]	-0.15*** (0.01)	0.25*** (0.008)	-48%***	3,860 [t]	-0.11*** (0.01)	0.25*** (0.008)	-44%***	3,392[ttt]
Within 180 days	-0.10*** (0.007)	0.17*** (0.004)	-59%***	10,484[ttt]	-0.13*** (0.01)	0.17*** (0.007)	-59%***	3,860 [t]	-0.10*** (0.01)	0.17*** (0.008)	-59%***	3,392[ttt]
Survival: Time-to-Death (in Days)												
Up to 2 years	61.75*** (15.39)	289.35*** (5.28)	21%***	10,388[ttt]	105.85*** (22.48)	255.71*** (6.77)	41%***	3,860 [t]	41.64*** (22.52)	297.66*** (10.50)	14%***	3,249[ttt]
Up to 1 year	47.69*** (11.15)	147.45*** (3.06)	34%***	10,388[ttt]	50.73*** (15.72)	136.38*** (3.86)	37%***	3,860 [t]	37.82** (15.09)	154.19*** (6.50)	25%**	3,249[ttt]
Up to 180 days	9.49 (7.15)	75.49*** (1.75)	0%	10,388[ttt]	6,76 (14.82)	71.97*** (2.26)	0%	4,215 [t]	8,65 (8.16)	75.26*** (3.51)	0%	3,249[ttt]

Table 3 (Cont'd)

	All CKD stages				Early CKD				Late CKD			
Survival Indicator:	ATEs (S.E.s)	PO Means (S.E.s)	ATEs/PO Means (%)	*n*, overid pass?	ATEs (S.E.s)	PO Means (S.E.s)	ATEs/PO Means (%)	*n*, overid pass?	ATEs (S.E.s)	PO Means (S.E.s)	ATEs/PO Means (%)	*n*, overid pass?
Time to Dialysis (days)												
Up to 2 years	80.37*** (18.38)	261.29*** (11.20)	31%***	10,388[ttt]	99.39** (45.35)	414.14*** (34.76)	24%**	3,860[ttt]	86.07*** (19.77)	221.21*** (12.57)	39%***	3,249[ttt]
Time to Higher CKD Stage												
Up to 1 year	25.70* (14.32)	293.35*** (7.73)	9%*	10,388[ttt]	17.90 (24.31)	335.65*** (12.81)	0%	3,860[ttt]	64.81*** (17.14)	218.38*** (10.19)	30%***	3249[ttt]
Time to Cardiovascular Disease Hospitalisation (days)												
Up to 2 years	41.34*** (14.29)	306.00*** (5.56)	14%***	10,484[ttt]	51.71** (27.62)	325.72*** (9.52)	16%**	3,860[ttt]	48.62*** (17.48)	276.82*** (9.25)	18%***	3,392[ttt]

Notes: As for Table 3.

CKD patients is similar (i.e., 24%). The largest estimated effect as a proportion of the POM is for PIRP patients with Late CKD, for whom dialysis is delayed by, on average, approximately 39% compared with their non-PIRP peers. In absolute terms the ATE for Early CKD is approximately 106 days, compared with approximately 42 days for the Late CKD results (i.e., the ATE for the former is approximately 254% of the latter). Once again, patients in the more advanced stages of CKD are expected, *ceteris paribus*, to be closer to potential dialysation in any event: but the effect of the intervention is nevertheless quite pronounced for patients with a diagnosis of Early CKD. This is also consistent with the better adherence results for these PIRP patients reported in the previous section.

The results on CKD disease progression in Table 4—which is observed by us only when a person is hospitalised and diagnosed with a higher-than-previous CKD stage—produce weak evidence that the PIRP program slows disease progression. Yet the results on time-to-hospitalisation for CVD show statistically significant delays of this outcome by the PIRP program of between 14 and 18 per cent.

6 CONCLUSIONS

The contemporary economic literature on CDMPs has started to take a nuanced approach to their effectiveness by concentrating on how their specific demand- or supply-side characteristics influence outcomes. Data limitations, though, have limited most of that literature to the effects of binary supply-side (e.g., nurse practitioner/GP) characteristics of an intervention, or binary demand-side (e.g., high-cost/low-cost patients) categorisations of outcomes. Often, these effects are observed over a relatively short time-frame.

This study focussed on a CDMP which adapts supply-side interventions to demand-side conditions (disease-staging) with a finer granularity and advances the contemporary literature in several novel and important ways. Using a dataset that spans seven years and contains detailed service use and health outcome data and an appropriate identification strategy, we were able to estimate the causal effects of the PIRP program on not only a host of adherence indicators, but also on fairly long-term health capital and health outcome indicators including disease progression, dialysis and deaths. Studying a progressive disease both with clearly-defined disease stages and a treatment regimen that is titrated by disease severity, also enabled us to produce nuanced estimates of the PIRP intervention's effectiveness. We were able to calculate dynamic, severity-specific, indicators of adherence to guidelines for pathology testing, specialist visits and pharmaceuticals, and to estimate the causal effect of the CDMP on these. Moreover, we were able to estimate its effect on health outcome indicators at different stages of the disease, thereby distinguishing the MHP of the program in early- and late-stage CKD patients. Our outcome indicators are observed over a long time-period—relative to the survival of this patient group—and included increases in disease severity and dialysis (indicators of quality-of-life) and survival/death (measures of the quantity of life).

The empirical results produced here show that this complex CDMP did improve many indicators of adherence and health outcomes and, in several instances, the estimated gains were considerable. A particularly important result is the finding that the intervention generally produced larger effects for patients with early-stage CKD than for those with late-stage CKD. This result is also of policy importance because, historically, CDMPs often have been implemented with the express view of targeting patients with advanced disease, based on the assumption that the resulting health product or cost savings will be greatest in that group than in patients with early-stage disease. On the contrary, our results provide support for the recommendation to invest heavily in monitoring activities and strategies to encourage adherence by chronic patients who are in their early stages in order to slow, significantly, the deterioration of their health capital.

In the foregoing respects, this study has gone further than other studies have been able to in respect of opening the "black box" of a CDMP and the outcomes it produces. While our approach provides new insights into the causal role of this intervention on both adherence and outcomes, its limitations should also be acknowledged. First, the evidence we produce on the program's effectiveness in respect of outcomes and on

adherence is produced in parallel, rather than as the output of a unified structural model. In this respect, we do not provide conclusive evidence that improved adherence is the driver of improved health outcomes under this program. Second, we do not address other important microeconomic questions: such as, which specific micro-economic features of the intervention influence patients and providers most strongly? These, and questions of how differences in the macroeconomic environment (e.g., the considerable vertical integration of health care and a single-payer health care system) may influence the effectiveness of this and other CDMPs, are important topics for future work.

REFERENCES

1. Raghupathi, W., Raghupathi, V. An empirical study of chronic diseases in the United States: A visual analytics approach to public health.
   Int J Environ Res Public Health. **15**, 431-454 (2018).
2. Brennan, P., Perola M., van Ommen G-J., Riboli E., on behalf of the European Cohort Consortium. Chronic disease research in Europe and the need for integrated population cohorts. Eur. J. Epidemiol. **32**, 741–749 (2017).
3. European Commission—Eurostat. Eurostat Statistics Explained: Health in the European Union—Facts and Figures, http://ec.europa.eu/eurostat/statistics-explained/index.php/Health_in_the_European_Union_%E2%80%93_facts_and_figures (2016). (Accessed 28 September 2019).
4. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. **390**, 1211-59 (2017).
5. Bleich, S.N., Sherrod, C., Chiang, A., Boyd, C., Wolff, J., DuGoff E., Salzberg C., Anderson K., Leff, B. & Anderson, G. (2015). Systematic review of programs treating high-need and high-cost people with multiple chronic diseases or disabilities in the United States, 2008–2014. *Preventing Chronic Disease: Public Health Research, Practice, and Policy*. **12**, http://dx.doi.org/10.5888/pcd12.150275 (2015). (Accessed 2 October 2019).
6. Simcoe, T., Catillon, M., Gertler, P. Who benefits most in disease management programs: Improving target efficiency. Health Econ. **28**, 189-203 (2019).
7. David, G., Smith-McLallen, A., Ukert, B. The effect of predictive analytics-driven interventions on healthcare utilization. J. Health Econ. **64**, 68-79 (2019).
8. Hussey, P.S., Schneider, E.C., Rudin, R.S., Fox, D.S., Lai, J., Pollack, C.E. Continuity and the costs of care for chronic disease. JAMA Int. Med., **174**, 742–48 (2014).
9. Amjad, H., Carmichael, D., Austin, A.M., Chang, C.H., Bynum, J.P. Continuity of care and health care utilization in older adults with dementia in fee-for-service Medicare. JAMA Int. Med. **176**, 1371–78 (2016).
10. Nyweide, D.J., Bynum, J.P.W. Relationship between continuity of ambulatory care and risk of emergency department episodes among older adults. Ann. Emerg. Med. **69**, 407-15e.3 (2017).
11. Imbens, G.W., Wooldridge, M.J. Recent developments in the econometrics of program evaluation. J. Econ. Lit. **47**, 5-86 (2009).
12. Imbens G.W., Rubin D.B. Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction. Cambridge (Mass.): Cambridge University Press (2015).
13. Muench, U., Guo, C. , Thomas C., Perloff, J. Medication adherence, costs, and ER visits of nurse practitioner and primary care physician patients: Evidence from three cohorts of Medicare beneficiaries. Health Serv. Res. **54**, 187-97 (2018).
14. David, G., Saynisch, P., Luster, S., Smith-McLalllen, A., Chawla, R. The impact of patient-centred Medical Homes on medication adherence? Health Econ. **27**, 1805-20 (2018).

15. David, G., Saynisch, P. & Smith-McLallen, A. The economics of patient-centred care. J. Health Econ. **59**, 60-77 (2018).
16. Soto-Gordoa, M., de Manuel, E., Fullaondo, A., Merino, M. Arrospide, A., Iguarta, J.I., Mar, J., on behalf of the CareWell Group. Impact of stratification on the effectiveness of a comprehensive patient-centered strategy for multimorbid patients. Health Serv. Res. **54**, 466-473.
17. World Health Organization. The Emilia-Romagna Region, Italy (Regions for Health Network Report), http://www.euro.who.int/__data/assets/pdf_file/0005/373388/rhn-emilia-romagna-eng.pdf?ua=1 (2018). (Accessed 2 October 2019).
18. European Commission. European Innovation Partnership on Active and Healthy Ageing, https://ec.europa.eu/eip/ageing/commitments-tracker/b3/effectiveness-evaluation-different-strategies-promoting-integrated-health.-0_en (2016). (Accessed 9 August 2019).
19. National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease: evalution, classification, and stratification. Am. J. Kidney Dis. **39**, S1-S266 (2002).
20. Rubin D.B. Estimating causal effects of treatments in randomized and nonrandomized studies. J. Edu. Psychol. **66**, 688-701 (1974).
21. Angrist, J.D. & Pischke, J-S. Mostly harmless econometrics: an empiricist's companion. Princeton: Princeton University Press (2009).
22. Imbens G.W., Rubin D.B. Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction. Cambridge (Mass.): Cambridge University Press (2015).
23. Wooldridge, J.M. Introductory Econometrics: A Modern Approach, 6th edn. Boston: Cengage (2016).
24. Rosenbaum, P.R., Rubin, D.B.. The central role of the propensity score in observational studies for causal effects. Biometrika. **70**, 41-55 (1983).
25. Lechner, M. Earnings and employment effects of continuous off-the-job training in East Germany after unification. J. Bus. Econ. Stat. **17**, 74–90 (1999).
26. Lechner, M. Program heterogeneity and propensity score matching: an application to the evaluation of active labor market policies. Rev. Econ. Stat. **84**, 205-220 (2002).
27. Barnow, B., Cain, G., Goldberger, A. Issues in the analysis of selectivity bias. Eval. Stud. **5**, 42–59 (1980).
28. Kaiser, B., Schmid, C. Does Physician Dispensing Increase Drug Expenditures? Empirical Evidence from Switzerland. Health Econ. **25**: 71-90 (2016).
29. Holm, S. A simple sequentially rejective multiple test procedure. Scand. J. Stat. **6**, 65-70 (1979).
30. Wooldridge, J.M. Econometric Analysis of Cross-Section and Panel Data. Cambridge (Mass.): MIT Press (2010).
31. Di Novi, C., Leporatti, L., Montefiori, M. Older patients and geographic barriers to pharmacy access: when non-adherence translates to an increased use of other components of health care. Health Econonomics (2020).
32. Gibertoni, D., Mandreoli, M., De Amicis, S., Cantarelli, C., Corradini, M., Caruso, F., Testa, F., Gasperoni, L., Orrico, C., Brancaleoni, F., Martelli, D., Angelini, M.L., Ferri, B., Marta Flachi, M., Iommi, M., Santoro A. L'accuratezza delle schede di dimissione ospedaliera nell'identificazione della malattia renale cronica e dei suoi stadi evolutivi. G. Ital. Nefrol. **36**, 65-76 (2017).
33. Martin, B., Wiley-Exley, E., Richards, S., Domino, M., Carey, T., Sleath, B. Contrasting measures of adherence with simple drug use, medication switching, and therapeutic duplication. Ann. Pharmacother. **43**, 36-44 (2009).
34. Sueta, C.A., Rodgers, J.E., Chang, P.P., Zhou, L., Thudium, E.M., Kucharska-Newton, A.M., Stearns, S.C. Medication adherence based on Part D claims for heart failure patients after hospitalization from the Atherosclerosis Risk in Communities Study. Am. J. Cardiol. **116**, 413-419 (2015).
35. Santoro, A., Gibertoni, D., Rucci, P., Mancini, E., Bonucchi, D., Buscaroli, A., Country, A., Hats, G., David, S., Gregorini, M.C., La Manna, G., Mosconi, G., Rigotti, A., Scorpions, R., Storari, V.,

Mandreoli, M. Challenges and results of PIRP project (Prevention of Renal Impairment Progressive) of the Emilia-Romagna. Ital. J. Nephrol. **5**, 1-9 (2019).
36. Charlson, M., Pompei, P., Ales, K., MacKenzie, C. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. J. Chronic Dis., **40**, 373–83 (1987).
37. Sundararajan, V., Henderson, T., Perry, C., Muggivan, A., Quan, H., Ghali, W.A. New ICD-10 version of the Charlson Comorbidity Index predicted in-hospital mortality, J. Clin. Epidemiol. **57**, 1288-1294 (2004).
38. Grossman, M. On the concept of health capital and the demand for health. J. Pol. Econ. **80**, 223-55 (1972).
39. Grossman M. The Human Capital Model, in Culyer A.J. and Newhouse J.P. (eds), The Handbook of Health Economics, Vol.1, Elsevier North Holland: Amsterdam (2000).
40. Imai, K., Ratkovic, M. Covariate balancing propensity score. J. R. Stat. Soc Series B. **76**, 243–263 (2014).