Better Off by Risk Adjustment? Socioeconomic Disparities in Care Utilization in Sweden Following a Payment Reform

Anell, Anders; Dackehag, Margareta; Dietrichson, Jens; Ellegård, Lina Maria; Kjellsson, Gustav

2022

Document Version:
Other version

Link to publication

Citation for published version (APA):
Anell, A., Dackehag, M., Dietrichson, J., Ellegård, L. M., & Kjellsson, G. (2022). *Better Off by Risk Adjustment? Socioeconomic Disparities in Care Utilization in Sweden Following a Payment Reform*. (Working Papers; No. 2022:15).

Total number of authors:
5

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Working Paper 2022:15

Department of Economics
School of Economics and Management

Better Off by Risk Adjustment? Socioeconomic Disparities in Care Utilization in Sweden Following a Payment Reform

Anders Anell
Margareta Dackehag
Jens Dietrichson
Lina Maria Ellegård
Gustav Kjellsson

August 2022
Revised: March 2024

LUND UNIVERSITY

Better Off by Risk Adjustment? Socioeconomic Disparities in Care Utilization in Sweden Following a Payment Reform

Anders Anell   Margareta Dackehag   Jens Dietrichson
Lina Maria Ellegård   Gustav Kjellsson

Author contact details: Anell: Department of Business Administration, Lund University. E-mail: anders.anell@fek.lu.se. Dackehag: School of Economics and Management, Lund University. E-mail: margareta.dackehag@ehl.lu.se. Dietrichson: VIVE - The Danish Center for Social Science Research. E-mail: jsd@vive.dk. Ellegård: Department of Economics, Lund University and Faculty of Business, Kristianstad University. E-mail: linamaria.ellegard@nek.lu.se. Kjellsson: Department of Economics, University of Gothenburg; Centre for Health Economics at University of Gothenburg (CHEGU), E-mail: gustav.kjellsson@economics.gu.se.

Funding statement: Funding by the Swedish Research Council for Health, Working Life and Welfare (FORTE, 2017-00877), Adlerbert Research Foundation, and Foundation for Economic Research in West Sweden is gratefully acknowledged.

Ethics approval statement: The analysis in this study has been approved by the regional ethics committee in Gothenburg, Sweden (Dnr 068-18, T777-18).

Data availability statement: The data in this study are confidential and cannot be shared publicly. The datasets may not be accessed without an approval from the Swedish Ethics Review Authority and an affiliation with Lund University or Gothenburg University. Since the dataset is constructed from official registers, researchers may also obtain the data from the sources, Statistics Sweden, the National Board of Health and Welfare, and the local health authority of Östergötland. These data are available for any researcher with an ethical approval from the Swedish Ethics Review Authority and subject to the regular confidentiality review of each register holder.

Conflict of interest disclosure: Not applicable.

Acknowledgments: The authors want to express their gratitude to Kim Rose Olsen, Anne Sophie Oxholm, Randall Ellis and participants at the 41th meeting of the Nordic Health Economists Study Group and the 2022 EuHEA conference for constructive comments. Warm thanks to Camilla Paananen for generously responding to questions about institutional details and sharing budget documents.

Better Off by Risk Adjustment? Socioeconomic Disparities in Care Utilization in Sweden Following a Payment Reform

Abstract

Reducing socioeconomic health inequalities is a key goal of most health systems. A challenge in this regard is that health care providers may have incentives to avoid or undertreat patients who are relatively costly to treat. Due to the socioeconomic gradient in health, individuals with low socioeconomic status (SES) are especially likely to be negatively affected by such attempts. To counter these incentives, payments are often risk adjusted based on patient characteristics. However, empirical evidence is lacking on how, or if, risk adjustment affects care utilization. We examine if a novel risk adjustment model in primary care affected socioeconomic differences in care utilization among individuals with a chronic condition. The new risk adjustment model implied that the capitation – the monthly reimbursement paid by the health authority to care providers for each enrolled patient – increased substantially for chronically ill low-SES patients. Yet, we do not find any robust evidence that their access to primary care improved relative to patients with high SES, and we find no effects on adverse health events (hospitalizations). These results suggest that the new risk adjustment model did not reduce existing health inequalities, indicating the need for more targeted incentives and interventions to reach low-SES groups.

Keywords: Socioeconomic health inequalities, Risk adjustment, Primary care, Health care utilization, Prospective payment, Incentives

JEL Classification: I11, I14, I18, L33, R50

1 Introduction

Socioeconomic health inequalities are ubiquitous. Whether measured by income, wealth, education, or occupation, individuals with lower socioeconomic status (SES) tend to live shorter lives, report worse self-assessed health, and suffer from more chronic conditions than individuals with high SES (e.g., Chetty et al., 2016; Mackenbach et al., 2008, 2018). Reducing health inequalities is one of the most important health policy objectives in many countries (e.g., Stabile and Thomson, 2014; Devaux, 2015).[1] In this study, we examine the effect of a policy aiming to reduce health inequalities by increasing the prospective payment for patients with high care need.

Prospective payment, i.e., *ex ante* payments intended to cover the expected care costs for an average patient during a particular time period or illness episode, gives incentives to select low-cost patients (cream-skimming), to avoid high-cost patients (dumping), and to undertreat patients (skimping) (Ellis, 1998, see e.g., Brown et al. (2014) and Werbeck et al. (2021) for empirical evidence of selection of profitable patients). Due to the generally lower health status and health literacy of low-SES patients (Paasche-Orlow et al., 2005), they are at elevated risk of suffering from dumping and skimping.

A common approach to avoid the negative consequences of prospective payment is to *risk adjust* payments using cost predictors such as diagnoses, demographic characteristics, and SES (Stabile and Thomson, 2014; Geruso and McGuire, 2016; Ellis et al., 2018).[2] Risk adjustment weakens the incentives for cream-skimming and dumping (Barros, 2003; McGuire et al., 2020). But since the payment remains prospective, there is no guarantee that an increase in the capitation of a given patient leads to increased care – i.e., less skimping – for the same patient. Providers, who can allocate the funds as they see fit, may prefer to retain the funds as profits, or to provide more care to other patients whose demand is more sensitive to signs of undertreatment. Thus, risk adjustment has theoretically ambiguous effects on socioeconomic inequalities in care utilization and health.

Given the extensive use, surprisingly few studies examine how providers react to risk adjustment (Brown et al., 2014; Geruso and McGuire, 2016; Layton, 2017). To our knowledge, there is no previous study of how risk adjustment affects SES-based inequalities. This paper contributes with such evidence from the context of primary care. As the first line of care, primary care providers guide patients in the health system, provide preventive services, treat common ailments and manage the treatment of several common chronic conditions (Rosano et al., 2013; Kringos et al., 2015). The quality and quantity of pri-

[1] Inequalities in health outcomes are to some extent mirrored by inequalities in access to and use of health care. Studies have for example found that low-SES patients make fewer specialist visits (Doorslaer et al., 2004; d'Uva and Jones, 2009; Cookson et al., 2016), use less preventive care (Devaux, 2015; Cookson et al., 2016), and consume less prescription pharmaceuticals (Nordin et al., 2013). The associations are sometimes smaller in primary care (Cookson et al., 2016), and seem to depend on the definition of SES and on how studies adjust for morbidity. For example, two Swedish studies find higher unconditional utilization among low-SES patients, and higher utilization among high-SES patients when conditioning on survey-based health measures (Agerholm et al., 2013; San Sebastian et al., 2017).

[2] For example, risk adjustment is used in the Medicare Advantage, Medicare Part D, the Health Insurance Marketplaces, and many state Medicaid Managed Care programs in the US, and in the health insurance markets of the Netherlands, Switzerland, Germany, Israel, and Belgium (Layton, 2017). The prospective payment to primary care providers is risk adjusted in for example Sweden (Anell et al., 2018), Denmark (Tange et al., 2020), and the UK (van de Ven and Ellis, 2000).

mary care can have important consequences for patients' health, including mortality effects (Starfield et al., 2005; Bailey and Goodman-Bacon, 2015; Ginja et al., 2022; Mora-García et al., 2024).

In our study setting – a mid-sized Swedish region (Östergötland) with around 450,000 inhabitants – primary care providers are mainly remunerated by capitation; i.e., they receive a fixed monthly sum for each enrolled patient. In 2014, the regional health care authority substantially changed the risk-adjustment formula in a way that implied a large increase in the capitation of low-SES patients relative to high-SES patients. Before the reform, the capitation was mainly adjusted for age. There was a SES-based adjustment factor, but it only affected the capitation of individuals living in the very poorest areas. The post-reform risk-adjustment formula instead incorporated a diagnosis-based morbidity adjustment (using the Johns Hopkins *Adjusted Clinical Groups* (ACG) system; Starfield et al., 1991), and a SES-adjuster based on individual characteristics such as low education, unemployment, and foreign background (the *Care Need Index*, CNI; Malmström et al., 1998; Sundquist et al., 2003). Due to the social gradient in health, both features of the novel risk-adjustment model disproportionately increased the capitation of low-SES patients, in particular those living outside the very poorest areas.[3] Indeed, the rationale for reforming the risk adjustment was to strengthen incentives to provide care for patients with high expected care need (Aldstedt, 2012; Zingmark, 2013)

The reform was announced in 2012, launched in 2014 and fully phased in by 2016. We use detailed register data on health care utilization from 2007-2017 to study how the reform affected individuals with a chronic condition in the ages 6-64. We use event study and difference-in-differences (DiD) approaches to compare the development of care utilization and adverse health events, i.e., hospitalizations, for low- and high SES individuals. Given the standard assumption of parallel trends, the estimates tell us how the reform affected the difference between individuals with high or low SES. This is a relevant estimand, given our aim to evaluate the impact of a redistribution of funds intended to reduce SES-based health inequalities. A comparative analysis using data from other regions, which did not simultaneously reform their payment systems, corroborates our results.

The reform implied substantially higher capitation payments for a majority of the low-SES individuals, but we find little to suggest that it improved their access to primary care. The DiD estimates on our main outcome measures, the probability and number of primary care physician visits, are small and negative. The estimates are similar across the dimensions of the SES index, but more negative for individuals whose provider was privately owned or projected to benefit substantially from the reform. We find a positive and statistically significant estimate on the probability of visiting a primary care nurse, but it is sensitive to specification changes and vanishes after the full phase-in of the model, when instead a negative estimate on the number of visits emerges. Moreover, a heterogeneity analysis indicates no changes for nurse visits among the patients with the very poorest health. For this highly prioritized

[3] For the individuals residing in the poorest areas, it is not clear that the payment increased, as the previous area-based SES adjustment in the old system may have exceeded the new ACG- and CNI-adjusted payment. Our results are robust to excluding the patients of the two PCCs located in the very poorest areas.

group, we thus find no traces of improved access to primary care.

We also consider outcomes outside the primary care sector. We do not find any statistically significant effects on the probability of a hospitalization or the number of hospital days. However, the probability of visiting the emergency department (ED), and the average morbidity risk score (the ACG weight) increased slightly more for low-SES group right after the reform. Given the small and statistically insignificant effects on hospitalizations, we do not think that the effects on ED visits and the ACG score reflects deteriorations of health. It is more plausible that the increase in ED visits reflects substitution from primary care, and that the increase in ACG reflects differences in diagnosis registration practices in the ED and primary care.

In sum, we find no evidence of substantially improved access to primary care and no indications of health improvements for a group that was intended to benefit from the new risk-adjustment model. Although there is no previous evidence on how risk adjustment affects SES-based inequalities, our results by and large resonate with the findings from related research. A descriptive study of three Swedish regions (including our study region) using the same SES-based index for risk adjustment did not find that the providers with more low-SES patients supplied more primary care visits (Anell et al., 2021a). However, relying on cross-sectional data, the study was not able to address the question of how the provision of care is affected by risk adjustment.

Of the few studies trying to study how risk adjustment affects the provision of care, the closest to our paper analyzes differentiated capitation in a laboratory experiment with medical students as subjects (Oxholm et al., 2019). The authors find that patients with similar needs receive more care if their capitation is above the average than if it is below the average. They also find that if the differentiation does not reflect patients' actual care need, there is a difference in the supply of care compared to under pure capitation. These findings suggest that physicians adjust their treatment choices in response to information about patient prioritization signalled by the payment differentiation. When the payment is aligned with physicians' prior information about care need, differentiated capitation does not alter treatment decisions relative to a pure capitation system.

Three studies from the US health insurance context use DiD strategies to examine provider responses to risk-adjusted capitation by comparing patients in Medicare Advantage with regular Medicare patients, for whom health insurers receive FFS payment. Of these, the closest to ours is Lissenden and Balkrishnan (2020), which compares the use of preventive services before and after the introduction of risk adjustment, using the change for FFS patients as comparator. They find that risk adjustment reduces pneumonia vaccination rates, but the effects on other preventive services are mixed and mostly not significant. Our study has a broader scope, as we consider the whole spectrum of primary care services rather than a narrow set of preventive services. Furthermore, compared to the health insurance context, the proposed mechanism behind the decrease in prevention (Eggleston et al., 2012) is weaker in our setting, where care providers are not accountable for secondary care costs. Geruso and Layton (2020), whose main focus is on diagnostic coding, find evidence of substantial upcoding as well as an increased probability of seeing a doctor in risk-adjusted Medicare Advantage. Brown et al. (2014) examine measures of beneficiary satisfaction and quality of care, and find little evidence of improvements.[4] A key difference between these studies and ours is that we consider responses to risk adjustment in a capitation setting, not in comparison to FFS.

As risk adjustment implies paying more for certain groups, our study also relates to the literature on changes of capitation rates. Duggan et al. (2016) and Cabral et al. (2018) show that increases of the level of capitation in Medicare primarily benefit health insurers, especially in less competitive markets. Duggan et al. (2016) additionally find increased entry and increased enrollment of traditional Medicare recipients, but no significant effects on patient satisfaction, self-reported health, or health care utilization. We provide evidence from outside the Medicare Advantage context. The health providers, who receive the capitation in our setting, cannot pass through payment increases to consumers in the form of lower co-payments. Another important distinction is that we study a policy change that did not necessarily increase the total payments for a given care provider. In our setting, the net effect on the payment to a given primary care practice depends on the risk profile of all enrolled patients.

Taken together, the results of our study and the previous related literature are consistent with the lack of strings attached to the prospective payment limiting the effects of risk adjustment. A policy implication is that other payment structures than risk-adjusted capitation may be preferable in relation to the goal of tackling socioeconomic health inequalities. Generally, service provision is often higher under fee-for-service (FFS) than under fixed payment schemes (capitation or salary) (Devlin and Sarma, 2008; Brekke et al., 2019; Cadena and Smith, 2022; Skovsgaard et al., 2023). There is experimental evidence to suggest that physicians are less likely to underserve high-need patients when paid by FFS (e.g., Hennig-Schmidt et al., 2011; Brosig-Koch et al., 2017). Furthermore, increased competition and patient choice are associated with a more pro-poor distributional change of visits to general practitioners (GPs) in regions that use mixed payments compared to pure capitation (Sveréus et al., 2018), and a recent study finds that an increase in the generosity of the FFS payment for low-SES individuals in Medicare led to increases in their care utilization (Cabral et al., 2021). This evidence suggests that mixed payment systems including a FFS component may be a more successful way to mitigate SES-based health inequalities than to rely on risk-adjusted capitation only.

In the next section, we discuss the theoretical implications of risk adjustment in more depth. Section 3 describes the primary care system and risk-adjustment models used in Sweden and Östergötland, and outlines our expectations regarding the effect of the reform. Section 4 presents our data sources and descriptive statistics of key variables. Section 5 describes our estimation strategy and Section 6 contains the results. Section 7 concludes.

[4]The focus of Brown et al. (2014)'s analysis, as well as the analysis in Newhouse et al. (2015), is risk selection and overpayment. The two articles come to partly opposing conclusions. Brown et al. (2014) find that risk adjustment, if anything, increased risk selection and overpayment, whereas Newhouse et al. (2015), using a longer panel, conclude that overpayment attributable to favorable risk selection decreased.

2 Theoretical consequences of risk-adjusting the capitation

In this section, we review the negative consequences of capitation that risk adjustment is intended to mitigate, and we describe why these consequences are especially pertinent to low-SES patients. We then emphasize that the effect of risk adjustment is ambiguous, and that it therefore may not reduce socioeconomic health inequalities.

In comparison to retrospective payments such as fee-for-service, prospective payment like capitation shifts the financial risk from the third-party payer to the health care provider. In other words, the marginal cost of treatment is borne by the provider. A downside of the strong incentives to economize on resources is that providers have incentives to select patients. As a result, they may overprovide care to profitable patients ("cream-skimming") and undertreat unprofitable patients (Newhouse, 1996). Undertreatment may materialize as explicit avoidance of unprofitable patients ("dumping") or as providers supplying too little care ("skimping"). Ellis (1998) shows that profit-maximizing providers in competitive markets skimp more on the care for high-cost patients.

Low-SES patients suffer higher risk than high-SES patients of being undertreated, for at least two reasons. First, their generally poorer health implies higher expected costs, which makes them less profitable (for a given level of capitation) and, as such, less likely to be cream-skimmed and more susceptible to dumping and skimping (c.f. Ellis, 1998). Second, with lower health literacy (Paasche-Orlow et al., 2005) and ability to seek, reach, and pay for care (Schwarz et al., 2022), low-SES patients are more likely to demand too little health care (Baicker et al., 2015). This weakens the exit and voice mechanisms that may otherwise limit skimping. Prospective payment may thus reinforce existing socioeconomic inequalities in health.

Risk adjusting the capitation, i.e., paying a larger (smaller) amount for patients with a high (low) predicted care need, is a common approach to mitigate the negative consequences of prospective payment (Stabile and Thomson, 2014; Geruso and McGuire, 2016; Ellis et al., 2018). Risk adjustment clearly weakens the incentives for cream-skimming of low-cost patients and dumping of high-cost patients (Eggleston, 2000). However, risk adjustment may not by itself change the incentives for skimping. Fundamentally, risk adjustment does not change the prospective nature of the payment, i.e., providers still bear the marginal cost of treatment. Whenever payments do not track costs well, i.e., the "fit" of the system is low (Geruso and McGuire, 2016), providers have weak incentives to provide additional treatments. Indeed, risk adjustment may even reinforce skimping on preventive services, as the providers' incentives to invest in prevention weaken when they can expect to be compensated for future cost increases (Eggleston et al., 2012).

Whether providers will respond to an increase of the capitation for high-cost patients by providing more care to the same patient group will depend on contextual features on the demand and supply side of the health care system. If patients are responsive to signs of undertreatment, and thus may respond to low quality via the exit mechanism, profit-oriented providers may respond to risk adjustment by increasing the care provided to high-cost patients as the competition for these patients increases (Ellis, 1998). If there is no credible exit threat from high-cost patients, the effect of risk adjustment depend on supply-side features. Providers who are profit-maximizing, or at least striving to avoid a budget deficit, may well find it optimal to allocate the additional resources obtained for enrolled high-cost patients to other patients who are more likely to exit, or to simply retain the money as profits. When it comes to (semi-)altruistic providers, it is plausible that increased capitation for high-cost patients would be used to increase the care provided to these patients, given their relatively high needs. However, if there are other patients with higher marginal benefit of care, providers may prefer to allocate more resources to them instead. Indeed, Barham and Milliken (2015) show that the existence of other, underserved, patients may lead semi-altruistic providers to *reduce* the care provided to their high-cost patients, following an increase in the capitation. In this regard, it should be recognized that risk adjustment implies a redistribution of funds from low-cost to high-cost patients. When positive and negative effects cancel out for a given provider, there is no impetus for a change in behavior (still assuming unresponsive patients).

Oxholm et al. (2019) argue that risk adjustment may affect treatment choices via the signal it sends to providers regarding the priority of different patients. I.e., the introduction of risk adjustment may be taken to indicate that high-need patients ought to be given even higher priority. The prerequisites for such a signalling mechanism is that providers are able to observe the patient characteristics that are used as risk-adjusters, and that the risk-adjustment model conflicts with providers' prior beliefs about prioritization.

Summing up, the impact of risk-adjusting the capitation on health care provision – and, consequently, on health inequalities – depends on detailed features of the risk-adjustment model and how information about the model trickles down to providers, as well as on characteristics of patients, providers and markets. In the next section, we describe the institutional background of our study and then discuss the plausible consequences of risk adjustment in our study setting.

3 Institutional background

Section 3.1 introduces the primary care system in Sweden and Östergötland. Section 3.2 describes the payment system and the studied reform. Section 3.3 details how the reform affected the capitation of low-SES patients relative to high-SES patients in our study population. Section 3.4 outlines our expectations regarding the effect on the socioeconomic difference in care utilization.

3.1 Primary care in Sweden and Östergötland

In the Swedish universal health insurance system, the responsibility for the financing and organization of health care resides with 21 regional health care authorities. Primary care is the first line of care and provides basic medical treatment, prevention, and rehabilitation to the whole population. Primary care providers also manage the care for patients with common chronic conditions such as hypertension, heart failure, type 2 diabetes, asthma, chronic obstructive pulmonary disease and dementia.

Providers are typically group practices – primary care centers (PCCs) – which are staffed by GPs (around 4-6), nurses and possibly other professions such as physiotherapists and cognitive therapists (Anell, 2015).[5] GPs and other staff are salaried employees. The regional health care authority contracts with public and private PCCs on equal terms. In our study region (Östergötland), there were between 43-46 PCCs during the study period.[6] In 2013, the nine private for-profit PCCs served one fifth of the population. The average number of patients was 8,800 for private and 10,500 for public PCCs.

Since September 2009, the region organizes primary care in a patient choice system with free entry. Providers that fulfil the (annually revised) accreditation criteria are allowed to establish a PCC anywhere in the region. Patients can visit any PCC, and they may also choose to enroll with a specific practice. PCCs are not allowed to reject patients who wish to enroll, and patients may switch PCCs as often as they like (Dietrichson et al., 2020). Most residents were enrolled with a PCC already before the choice reform, but the PCCs did not have to accept new patients and thus choice was more limited.

There were two other notable changes during the study period. First, in 2014, the municipalities overtook the responsibility for home visits by nurses to elderly persons (65+) from the region, i.e., some district nurses became employed by the municipalities instead. This reform had no direct implications for the GPs. We handle this confounding policy by excluding elderly individuals from the study population. Second, the patient co-payment for GP visits was raised from SEK 150 to SEK 200 (from 100 to 200 for nurse visits) in 2017.[7] Earlier Swedish studies show that patients, and in particular low-income groups, respond to increasing fees by making fewer visits (Nilsson and Paul, 2018; Johansson et al., 2019). However, our conclusions remain unchanged if we exclude 2017 from the analysis.

3.2 Payment system and reform

Throughout our study period (2007-2017), the payment to PCCs predominantly consisted of capitation for enrolled patients (75-85% of payment), complemented by additional payments to account for structural features and some pay-for-performance. Additional details of the payment schemes are described in Appendix A.[8]

[5] Outside the regional organization, there is a small group of physicians running private solo practices funded by the central government budget. (This arrangement is a remnant of a previous system.) As these are not funded by the region, they are not pertinent to the present study..

[6] Three private PCCs opened in October 2009, November 2010, and October 2012. Two public PCCs closed in July 2011 and one in April 2016. There was a net increase (+1 PCC) in the largest city and a net loss (-1 PCC) in the second largest city.

[7] These fees applied when patients' visited the PCC they were enrolled at. To discourage patients from visiting another PCCs than the one that they were enrolled at, the visit fee was SEK 500 for visits at other PCCs. In 2017, the exchange rate of SEK was 9.6 to the Euro, 8.5 to the USD, and 11.0 to the GBP (Riksbanken, 2021). National regulations cap the total fee paid annually; in 2010-2011 the cap was set at SEK 900 and in 2012-2017 at SEK 1,100. Children below 20 were exempt from fees.

[8] All information in this section is, unless otherwise mentioned, taken from the region's primary care budgets, or, for 2010-2017, the yearly terms of accreditation for PCCs or from protocols from regional board meetings, which are available on request from the authors (in Swedish).

The capitation has always been risk-adjusted, but the risk adjusters have varied over time as described by Table 1. Before 2014, the capitation was adjusted by age; the capitation for prescription drugs was also adjusted by gender. Additionally, the capitation was higher for elderly individuals (75+) living in remote areas and higher for residents in the very poorest areas. For individuals aged 20-44 years (45-65), this SES compensation corresponded to around 150% (100%) of the base capitation. Notably, the SES compensation affected PCCs very unequally: In 2010, it accounted for less than 3 percent of payment for 27 of the 43 PCCs, but for more than 9 percent for the four PCCs located in highly deprived areas.

Table 1: Adjustments of the capitation

Year	**Age**	**Elderly**	**Gender**	**Morbidity**	**SES**
Pre 2014	Yes	In remote areas	(Drugs)		Poor area
2014-2015	Yes		Yes	ACG	CNI
2016-2017	Yes	Yes	Yes	ACG	CNI

With the 2014 payment reform, Östergötland moved to a risk-adjustment model in which 90% of the base capitation was adjusted for the expected care costs given the patient's diagnoses, age and gender using the Johns Hopkins Adjusted Clinical Groups (ACG) system (which is widely used internationally Handel et al., 2015). The motivation for the reform was that policymakers feared that the previous model induced cream-skimming and that PCCs with many high-cost patients did not receive fair compensation for their responsibilities. Preliminary simulations showed that the fit of the model would improve considerably when using ACG.

While the ACG does not depend directly on socioeconomic status, the socioeconomic gradient in health implies that low-SES patients on average have higher risk scores, a feature that the region was well aware of. Indeed, channeling funds towards the low-SES group was an intended purpose of introducing the ACG adjustment (Aldstedt, 2012). However, since patients with low SES may require more care for a given level of illness, the region added a SES-based adjustment on top of the ACG (Zingmark, 2013). Studies from other Swedish regions indicate that low-SES patients receive less primary care in relation to their need than high-SES patients (Agerholm et al., 2013; San Sebastian et al., 2017; Gustafsson et al., 2024).

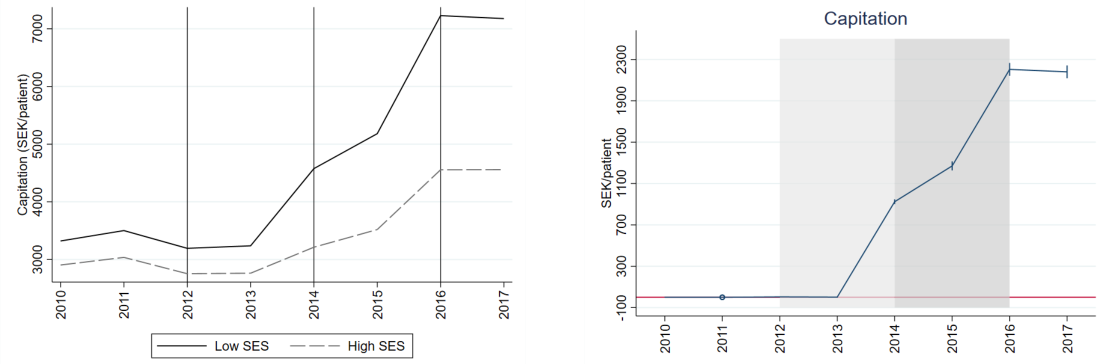
Östergötland replaced their previous area-based SES compensation by an individual-level adjustment based on the Care Need Index (CNI) (Anell et al., 2018), an index reflecting the relative workload (as judged by Swedish physicians) associated with seven patient characteristics (Malmström et al., 1998; Sundquist et al., 2003): Being under five years of age (weight = 3.23); being born in Africa, Asia, South America, or in a southern or eastern European country that is not a member of the European Union (EU; 5.72); being over 65 years and living alone (6.15); being a single parent with children under 17 years (4.19); being 1 year or older and recently having moved to the area (4.19); being 16-64 years and unemployed (5.13); and being 25–64 years and having at most nine years of schooling (3.97). CNI contains both SES and demographic characteristics, but it is highly correlated with low income (Anell et al., 2021a) and in the setting for our analysis it is a pure SES-measure due to the definition of our study population.[9]

The decision to introduce ACG was taken by the regional health care board on May 8, 2012, and mentioned alongside CNI in the terms of accreditation for 2013, which were approved on May 30, 2012 (Hälso- och sjukvårdsnämnden, 2012). The new payment model was phased in during 2014-2015 and fully implemented by 2016.

3.3 How did the reform affect the capitation of low-SES patients?

Figure 1(a) shows the average capitation of our study population, which consists of individuals aged 6-64 with a chronic condition (see section 4.2), by year and SES in 2010-2017 (we lack data on the age weights for 2007-2009). The vertical lines indicate the start of the announcement period (2012-2013), the phase-in period (2014-2015), and the post period (2016-2017), respectively.

The figure shows that the average capitation was slightly higher in the low-SES group already before the reform, reflecting differences in demographics and the area-based SES compensation of the old capitation system. However, this pre-existing difference is dwarfed by the divergence in 2014, when the new system was rolled out. Although both groups benefited from the reform on average, the increase was much larger for the low-SES group. This is confirmed in Figure 1(b), which shows estimates from an event-study model (see Eq. (1) in section 5.2). These results show that the low-SES patients in our study population became more profitable due to the reform, both in absolute terms and relative to high-SES patients.



(a) Levels

(b) Event study estimates

Figure 1: Capitation

There are two caveats to this general picture. First, for individuals living in the poorest areas, the loss of the earlier area-based payment may not have compensated for the ACG and CNI payment. Indeed, the total capitation payment fell substantially in 2014 for the two PCCs that used to receive the very

[9]The study population excludes children and elderly individuals, and we do not classify people who moved within the region as having low-SES. The region omitted this factor from 2016 onwards, as it was not viewed as a relevant measure of SES.

highest compensation from the SES adjustment in the old payment system. From 2016 onwards, the region compensated these two PCCs with extra lump sum grants due to their SES burden. As we show in section 6.2.1, our main results are robust to excluding patients at these PCCs from the analysis.

Second, the new risk-adjustment model did not increase the capitation for all low-SES individuals all years in our study population. For instance, for the quartile of the low-SES patients that benefited the least from the reform, the capitation in 2013 would have been reduced by at least 9 percent if it had been calculated according to the 2014 rules. The corresponding number among high-SES patients is much larger: the capitation for the most disadvantaged quartile would have been less than 30% of the capitation based on the 2013 rules. However, it is important to note that very few individuals were hurt by the reform in every year. The size of the capitation post reform depends heavily on the ACG, and almost everyone in our study population had an above-average ACG in at least one study year. Importantly, our results are not driven by the few low-SES individuals who permanently became less profitable due to the reform (see section 6.2.1).

3.4 Expected consequences of the reform on care utilization

Given the institutional background, we are now in a position to discuss the plausibility and expected outcomes of the mechanisms described in the theoretical framework in section 2. For clarity, we discuss the expected effects of each mechanism holding everything else constant, even though our empirical approach does not allow us to study each mechanism in isolation.

In a setting with salaried GPs, a fundamental question is whether providers will react to financial incentives. We expect the PCCs to respond to incentives. The private PCCs are for-profit firms, and for public PCCs, showing a decent surplus or at least breaking even is an important performance measure for the managers (Vengberg et al., 2021). Empirical studies from Sweden also support the notion that public as well as private PCCs respond to financial incentives (Anell et al., 2018; Ellegård et al., 2018; Dackehag and Ellegård, 2019; Ellegård, 2020; Vengberg et al., 2021).

The question then is what incentives the reform implied. As noted in section 2, risk-adjustment does not change the fact that prospective payment gives providers autonomy over the allocation of funds across patients. The average increase in capitation suggests that the reform made low-SES patients more attractive for cream-skimming, and less susceptible to skimping or dumping. If PCCs compete for patients, then the equilibrium level of care offered to low-SES patients should be higher, the more profitable they are. The margin for outright dumping is limited in our setting, where there is a universal right to enroll at any PCC and almost everyone is enrolled. Nonetheless, the increased profitability may induce intensified competition for the high-costs patients. Thus, all else equal, the competition mechanism suggests that low-SES patients would be expected to use more care relative to high-SES patients after the reform.

A precondition for the competition mechanism to work is that low-SES patients are responsive to undertreatment – i.e., that there is an exit threat. Studies from Sweden, Norway, the UK, and the US report that patients with poor health or low-SES have relatively low demand responsiveness to quality (Tay, 2003; Biørn and Godager, 2010; Gutacker et al., 2016; Santos et al., 2017; Anell et al., 2021b), suggesting the exit threat may be weak. Furthermore, within our study population of patients with chronic conditions, the reform increased the average profitability of high-SES patients too (as shown in Figure 1). These observations suggests that the reform may not have affected the SES difference in care utilization, and that the low SES patients' care utilization might even have decreased relative to that of high-SES patients.

Evidence from a wide range of settings (including other Nordic countries) suggests that GPs are (semi-)altruistic (Yordanov et al., 2023; Galizzi et al., 2023). Altruistic providers may react to increased capitation by providing more care to other high-need groups (Barham and Milliken, 2015). In our setting, young children and elderly patients are examples of such groups. If this mechanism is at play, we expect to observe decreasing levels of care for both the low- and high-SES individuals in our study population (since both groups experienced increases of the average capitation), leading to an ambiguous effect on the socioeconomic difference in care utilization in the study population. Furthermore, the altruism mechanism should mainly operate for providers whose total revenues changed substantially after the reform. For providers whose total revenues changed little following the reform, we do not expect the altruism mechanism to affect the care provided to any part of our study population.

We do not believe that the reform would aggravate skimping by diluting the incentives for prevention (Eggleston, 2005). The largest savings potential from prevention would arise outside the primary care sector, i.e., from avoided hospitalizations. It is unlikely that PCCs would take such savings into account.

With regards to a signalling effect (Oxholm et al., 2019) , we note that concerns about undertreatment of low-SES patients was an explicit motivation for reforming the risk-adjustment model. Thus, if anything, the reform signalled that the providers were currently not providing enough care to low-SES patients. If this mechanism is important, we would expect an increase in the care utilization for low-SES patients relative to high-SES patients. In particular, we expect larger positive effects on patients with readily observed characteristics that are associated with a higher care need. In the Swedish setting, having immigrant background from outside the EU is fairly observable. In contrast, it is less easy to predict if someone has only primary or upper secondary schooling.

Although our setting constrains the set of mechanisms through which risk-adjustment may affect undertreatment, the above discussion shows that the theoretical effect on the SES difference in care utilization remains ambiguous. In the next section, we describe the data we use to obtain an empirical estimate of the effect.

4 Data

4.1 Data sources

We use Östergötland's register over enrollments at PCCs to define our study population. This register is available from 2008 and covers all residents, even if they have not chosen to enroll with a specific provider. Data on primary care utilization, emergency department (ED) visits, and diagnoses come from the regional care register. We use the national patient register from the National Board of Health and Welfare to define variables related to inpatient care (overnight hospital stays), and individual background data from national registers held by Statistics Sweden to define CNI weights. Finally, we source information about the projected budget impact from regional health care administrators.

4.2 Study population

We define our main study population as follows. We start from a list of the total population in Östergötland that was enrolled at a PCC on January 1 2013, i.e., one year before the new payment system came into effect (433,312 individuals). We then apply the following inclusion criteria:

The individual must reside in the region throughout the pre-period (2007-2011), and be born before 2002 and after 1953. The lower age limit is justified by the fact that all children below 5 years of age have a high CNI per definition, so they lack a comparison group. The rationale for the upper age limit is that individuals who turned 65 during our study period might be affected by the transfer of responsibility for home health nursing from the region to the municipalities. As the main reason why elderly individuals have a high CNI is that they are living by themselves, the home care reform likely had a differential impact on low- and high-SES elderly individuals. Single-living elderly are often widows, i.e., relatively old, and many of the most frail persons belong to this category. Thus, older individuals with high CNI were likely more affected by the re-organization of home care than older individuals with low CNI.

Finally, we restrict the study population to individuals who had a chronic condition diagnosis recorded in at least one of the pre-reform years.[10] The reason is conceptual: the objective of the study is to examine if risk adjustment of the capitation can reduce the skimping and dumping problems that, in a pure capitation system, would disproportionately affect low-SES individuals in poor health. How the reform affected low-SES individuals in good health – whose capitation may well have been reduced due to the reform we study – is another question. Although it would be interesting to study this question, data presented in Online Appendix J suggests that we cannot convincingly do so with our empirical strategy.

In summary, our main study population is a cohort of adolescents and young to middle-aged adults, who resided in the study region in the whole pre-period and had a chronic condition before the reform. The final study population includes 92,863 individuals.

[10] We use the chronic condition count produced by the Johns Hopkins ACG (R) System v 11.1 to identify chronic conditions. See Online Appendix B.

4.3 Variable definitions

4.3.1 SES definition

Our estimations contrast individuals with low SES to a comparison group of individuals with high SES. We use the CNI to define individuals' SES. Specifically, we define an indicator variable, *High CNI*, that equals 1 for individuals in the study population who possessed any of the characteristics associated with a CNI weight above 1. Given the definition of our study population, this implies that we define as low-SES individuals people who were born in Africa, Asia, South America or in non-EU European countries, are single parents, lack more than primary education after reaching 25 years of age, and/or are unemployed.

As mentioned, children below five and elderly living alone would also count as low-SES/high-CNI individuals, but they are excluded from our analyses. The original CNI also assigns above-1 weight to individuals who changed address (moving within or to the region) during the past year, but we classify such individuals as belonging to the comparison group as we do not view it as an informative proxy for SES. Notably, the region removed this component from the calculation of payment in 2016 because it only benefited PCCs in locations with high inward mobility, i.e., close to university student housing areas (Aldstedt et al., 2015). The descriptive statistics for the comparison group are almost identical when including and excluding this small group of people. In the following, we use the terms low-SES/high-CNI interchangeably (and similarly for high-SES/low-CNI).

We define individuals' SES based on their CNI on December 31, 2011, which is our most recent pre-announcement data point. To avoid post-treatment bias, we do not update the treatment definition over time. (The results are robust to a time-variant treatment definition, see section 6.1.)

4.3.2 Outcome variables

Our primary outcome variables are the probability of a visit and the number of visits to a *physician* at a PCC. As nurses perform many services at PCCs, we also construct the corresponding variables for *nurse* visits. We define these variables at a quarterly level. As our data do not include information on the type of services provided, we are not able to study preventive activities *per se*.

Our set of secondary outcome variables includes outcomes that may be affected by changes in primary care utilization, either because the outcomes are directly affected by the volume and quality of primary care, or because they may substitute for primary care. We examine the individual's morbidity weight from the ACG system (the ACG weight).[11] Notably, the introduction of ACG-based payment increases providers' incentives to more carefully register diagnoses (van de Ven and Ellis, 2000), which

[11] We used the official ACG software, version 11.2.1, to compute ACG weights. Our measure of the ACG weight does not correspond exactly to the weights used by the region for the payment. The region computes ACG weights based on diagnoses registered over 18 months. We computed the ACG weight including only diagnoses registered during a calendar year, for convenience given the structure of our data (annual). Furthermore, the ACG weights used in the regional payment system only includes diagnoses set in primary care. We included all diagnoses in our computation of the ACG weights, to get a more comprehensive measure of individuals' health status.

might have heterogeneous effects on low- and high-SES patients.

We also study measures of secondary care utilization and health: The probability of a visit to a hospital emergency department (ED), the probability of being hospitalized (inpatient stay), the number of days spent in hospital, and the probability of a hospitalization with a so-called ambulatory care sensitive condition (ACSC).[12] ACSC hospitalizations are often referred to as *avoidable* given appropriate prevention and primary care. This is a measure of adverse health events and indicative of low primary care quality. The secondary outcome measures are computed on an annual basis; the ACG weight because a sufficient time period is needed to collect information on diagnoses, the hospital measures because they are rare events.

4.4 Descriptive statistics

Panel A of Table 2 shows descriptive statistics for indicators of SES status for the two groups contrasted in our analysis. The first row shows that on average, individuals in the low-SES group (high CNI) group belonged to the high-CNI group during 84% of the study period. This suggests that our approach of using a time-invariant measure of SES is reasonable. The second row shows the mean CNI weight in 2011, and the subsequent rows show the proportion of individuals possessing each of the CNI characteristics in our low-SES definition. By definition, the average CNI weight is higher in the low-SES group. (The reason why the average CNI is larger than zero in the high-SES group is that we classified individuals whose CNI is high *just because they moved* as belonging to the high-SES group.)

We note from the table that low education, unemployment and immigrant status are the most common reasons for being categorized as having low SES. The labor income is considerably lower in the low-SES group, confirming that our analysis contrasts groups with substantially different SES.

Panel B of Table 2 shows descriptive statistics of variables related to health care utilization and health status in 2011. Despite that all individuals in our study population had a chronic condition, we see that individuals in the high CNI group on average visited physicians and nurses more, in primary care (PC) as well as in the whole health care sector (HC), than individuals in the low CNI group. They were more likely to visit the ED, slightly more likely to be hospitalized, and stayed more nights in hospital on average. The ACG weight is higher in the low-SES group, which confirms that they have relatively low health status even conditional on having a chronic condition. Another indication of the lower health status in the low-SES group is that 16% received disability pension, compared to 9% of the comparison group. In Appendix C we show that, conditional on the ACG weight, there is little difference in health care utilization between the SES groups.[13]

[12] We use the definition of ACSC developed by the National Board of Health and Welfare. The set of ACSCs include the following chronic conditions: anemia, asthma, diabetes, heart failure, hypertension, COPD, angina, the following acute conditions: bleeding ulcer, diarrhea, epileptic seizures, inflammatory diseases in the female pelvic organs, pyelitis, and ear, nose and throat infections, and the following conditions of special relevance to elderly patients: cardiac arrhythmia, influenza, pneumonia, and urinary tract infections.

[13] Notably, the ACG score is not a perfect measure of health, as it depends on the individual's health care consumption. If,

Table 2: Descriptive statistics

Panel A: CNI components and income	High CNI (treated)			Low CNI (comparison)		
	N	Mean	SD	N	Mean	SD
High CNI (within ind.)	26663	0.84	0.25	66200	0.07	0.14
CNI weight (t=2011)	26663	5.73	2.86	66200	0.23	0.93
Single parent	26663	0.19	0.39	66200	0.00	0.00
Foreign	26663	0.31	0.46	66200	0.00	0.00
Low education	26663	0.41	0.49	66200	0.00	0.00
Unemployed	26663	0.36	0.48	66200	0.00	0.00
Moved	26663	0.07	0.25	66200	0.06	0.23
Labour income (SEK)	26663	139196	150756	66200	255873	190334

Panel B: Health and care variables	High CNI (treated)			Low CNI (comparison)		
	N	Mean	SD	N	Mean	SD
GP visits (PCC)	26663	1.56	1.97	66200	1.20	1.55
Nurse visits (PCC)	26663	1.46	4.74	66200	1.18	3.47
Phys visits (HC)	26663	3.26	3.91	66200	2.71	3.55
Any ED visit	26663	0.23	0.42	66200	0.17	0.38
Any inpatient stay	26663	0.11	0.31	66200	0.09	0.28
Any planned inpatient	26663	0.03	0.18	66200	0.03	0.18
Any acute inpatient	26663	0.08	0.28	66200	0.06	0.24
Hospital days	26663	0.97	7.35	66200	0.72	6.03
ACSC hospitalisations	26663	0.01	0.10	66200	0.01	0.08
ACG weight	26663	1.53	1.35	66200	1.27	1.17
Disability benefit	26663	0.16	0.36	66200	0.09	0.29

Note: **Panel A** shows the proportion of sample years with high CNI (within individual variation), the CNI weight and components of the CNI in 2011, and annual labour income. **Panel B** shows descriptive statistics for the health related variables. Mean values (proportions) presented separately for individuals with high CNI (the treatment group) and low CNI (the comparison group) as of December 2011. High value of the CNI indicates low socioeconomic status. The comparison group includes individuals who have a strictly positive CNI due solely to having moved within the region. All variables are measured on an annual basis using data for 2011. The summary statistics are weighted to balance the treatment and comparison groups in terms of gender and birth year. PCC is primary care, HC = all health care, ACG weight is the individual's ACG weight calculated from diagnoses set in 2011, ED = emergency department, ACSC = ambulatory care sensitive conditions.

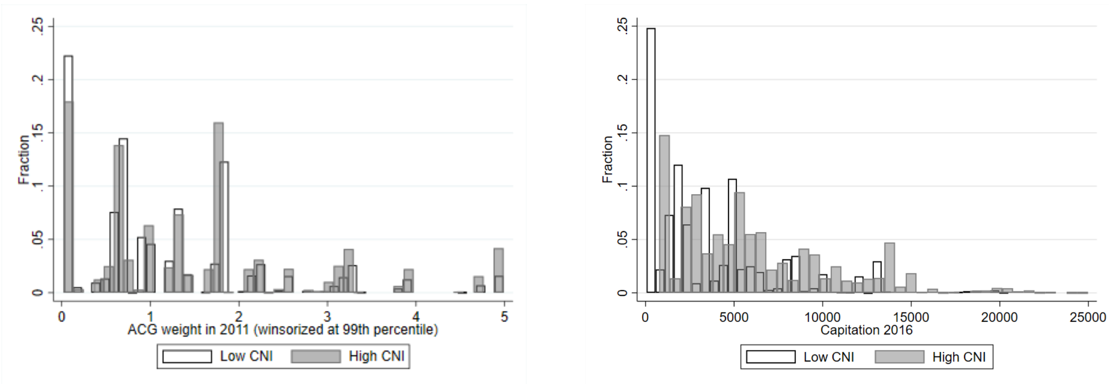
Note: **Panel A**: The first row shows descriptive statistics for the proportion of the years in the sample that individuals are classified as having high CNI (within individual variation). The subsequent rows show descriptives for the CNI weight, the CNI components, and annual labor income in 2011. **Panel B** shows descriptive statistics for the health related variables. PCC is primary care, HC = all health care, ACG weight is the individual's ACG weight calculated from diagnoses set in 2011, ED = emergency department, ACSC = Ambulatory Care Sensitive Conditions. **Both panels:** The descriptives are presented separately for individuals with high CNI (i.e., low SES) and low CNI (high SES) as of December 2011. Note that we have classified individuals who would in effect have a strictly positive CNI due solely to having moved within the region as belonging to the low-CNI group (see main text). All variables are measured on an annual basis using data for 2011. The summary statistics are weighted to balance the treatment and comparison groups in terms of gender and birth year.

The average ACG weights exceed one in both groups, which means that their health care costs were above the regional population average. This is reasonable, since the study population only includes individuals with a chronic condition. To gain further insights on the morbidity in the low- and high-SES groups, Figure 2 (a) shows the distribution of ACG weights in 2011, by group. Note that the ACG variable is highly skewed with a small number of outliers. To make the figure interpretable, the weights are winsorized at the 99th percentile. I.e., individuals with ACG above the 99th percentile have been assigned the weight of the 99th percentile. Although it is clear that low-SES individuals are overrepresented in the

all else equal, high SES individuals are more prone to seek, or get access to, care for a given condition (e.g., Agerholm et al., 2013; Angerer et al., 2019; Schwarz et al., 2022), the ACG weight of low-SES individuals will overestimate their health relative to high-SES individuals and the comparison of health care utilization at a given ACG weight underestimates actual SES differences in health care utilization.

higher ACG risk groups, there is a considerable overlap between the distributions. Consequently, there is also some overlap in the distributions of the capitation after the reform (when it became partly based on ACG), which is illustrated by Figure 2(b). Yet, the socioeconomic gradient in the capitation is clear – 25 percent of the high-SES group had lower capitation than the lowest capitation observed in the low-SES group (around SEK 800).

The proportion of individuals classified in the very lowest ACG is remarkably high for both groups (18-20%), given that the study population only includes individuals with a chronic condition. The high proportion reflects that we only include 12 months of data when calculating the ACG. For idiosyncratic reasons on both the supply and demand side, patients with chronic conditions do not always attend checkups within a 12-month period. As we show in section 6.2.1, almost all individuals in the study population are classified in ACG groups corresponding to above-average health costs (i.e., > 1) at least once during the study period.



(a) ACG distribution (pre-reform) (b) Capitation distribution (post reform)

Figure 2: ACG and capitation distributions by SES (high CNI = low SES)

5 Empirical strategy

This section first presents our identification strategy and discusses threats to this strategy (section 5.1). We then describe how we implement the strategy in our estimations (section 5.2).

5.1 Identification

We compare the high and low SES groups in an event-study and DiD framework. As in any DiD analysis, the extent to which the estimates can be interpreted as being driven by the reform depends on the plausibility of a parallel trends assumption (e.g., Abadie, 2005). Given that this assumption holds, the estimates tell us how the reform of the capitation model affected the difference in outcomes between individuals with high or low SES. As the goal of our analysis is to evaluate a redistribution of funds initiated with an intention to affect SES-based health inequalities, this is a relevant estimand.

Note, however, that our analysis does not capture the total effect of the reform. Since the reform affected the capitation of all patients, the outcomes in the high-SES group do not describe the counterfactual development of outcomes for the low-SES group if the payment model had been unchanged. Furthermore, our analysis does not answer the question of how the outcomes in the low-SES group would have evolved, had the health authority increased the total payment and channelled those extra funds to the low-SES group.

The major threat to the parallel trends assumption in our case is the possibility of diverging health trends. In particular, the generally lower health of low-SES individuals may deteriorate more quickly. This would imply increases of both primary and secondary care utilization over time relative to the utilization of high-SES individuals, irrespective of the reform. As we restrict the sample to individuals with a chronic condition, this is a less likely source of bias than it would have been had we studied the whole population.[14] Individuals with chronic illnesses are receiving care and monitoring from the health care system, which make it more likely that the small difference in the level of care utilization between the treatment and comparison group is kept in an equilibrium, unless upset by an external shock such as, e.g., a new risk-adjustment model.

To further mitigate the scope for differential health trends, we balance the high- and low SES groups in terms of age and gender by including age- and gender specific weights obtained by Coarsened Exact Matching (CEM; Iacus et al., 2011) in our regressions. We also examine the robustness of our results to including group-specific linear trends, implement a synthetic DiD (Arkhangelsky et al., 2021), and we compare the development in Östergötland to three other regions that did not implement reforms of the payment system (section 6.2.3). None of these analyses suggest that diverging health trends is a problem in our sample.

5.2 Estimation

We estimate flexible event-study specifications for all outcomes. We interact a vector of time-fixed effects with an indicator variable for belonging to the low-SES/high-CNI group. In our models of primary care utilization, we have 44 quarters of data (2007-2017) and thus the event-study specification for outcome y becomes:

$$y_{iq} = \sum_{j=1}^{j=44} \gamma_j \mathbf{I}(\text{High CNI}) \times \mathbf{I}(j) + \lambda_q + \mu_i + \varepsilon_{iq} \tag{1}$$

where y_{iq} is the value of the dependent variable for individual i in quarter q, $\mathbf{I}(\text{High CNI}) \times \mathbf{I}(j)$ is the interaction between the low-SES indicator and the FE for quarter j, λ_q is the vector of quarter FEs, μ_i is the vector of individual FEs and ε_{iq} is an error term.

[14] An alternative approach to reduce the scope of this problem would be to study a population defined by age instead of a cohort. In a sensitivity analysis, we consider such a study population in which individuals enter the sample when they turn six, and leave the sample when they turn 65. The population is redefined yearly using the population enrolled at a PCC in the region on January 1. Because the region started to record the data on PCC enrollment in 2008, we can only analyse this time-varying study population in 2008-2017.

We cluster standard errors by the individual. As the capitation differs across individuals, this is the level of treatment assignment. Furthermore, there are both low and high-SES individuals enrolled at every PCC in the region, and the treatment is not received in groups. That is, common reasons to cluster at higher levels are not present (Lohr et al., 2014; Weiss et al., 2016; Abadie et al., 2023). We present the estimated vector of γ_qs in graphical format, using the last quarter of 2011 as the reference category.

To gain precision, we also estimate more restrictive DiD models contrasting the whole pre-period (2007-11) with the announcement (2012-13), phase-in (2014-15) and final post reform (2016-17) periods. We build these aggregate estimates from the following event-study-like specification:

$$y_{iq} = \sum_{j=21}^{j=44} \gamma_j \mathbf{I}(\text{High CNI}) \times \mathbf{I}(j) + \eta \mathbf{I}(\text{High CNI}) \times q + \lambda_q + \mu_i + \varepsilon_{iq} \tag{2}$$

The specification is equivalent to (1), except that it restricts all pre-period estimates to be 0 and allows for a separate linear trend for the high-CNI group ($\eta \mathbf{I}(\text{High CNI}) \times q$); although in the baseline specifications, we restrict the coefficient η to zero (thus assuming similar trends). We use this approach because it allows us to include a group-specific linear time-trend for the treatment group by extrapolating the potentially differential trend from the pre-period (η; see Bilinski and Hatfield, 2018). When we do include such a trend, the fully dynamic specification in eq. (2) helps us ensure that the group-specific trend is estimated using only variation in the pre-period as recommended by Wolfers (2006) and Lee and Solon (2011).

To obtain DiD estimates from the quarterly estimates, we compute averages of the γ_qs from Eq. (2) over three separate periods: i) the announcement period (2012-2013, i.e., $q \in (21,28)$), the phase-in period (2014-2015, i.e., $q \in (29,36)$), and the period after the new system was fully phased in (2016-2017, i.e., $q \in (37,44)$). Since we are not in a staggered DiD setting, this approach yields exactly the same point estimates as a standard DID specification using *post* × *treatment* dummies (Goodman-Bacon, 2021). We use the delta method to estimate standard errors for the estimated DiDs (clustered by individual).

We also estimate event-study and DiD models for our secondary outcomes (hospitalizations etc). Because these are rare events, we use annual instead of quarterly data in these estimations, and we use 2011 as the reference year.

6 Results

Section 6.1 presents our main results: the effects of the reform on primary care utilization. In section 6.2, we extend the analysis by exploring the heterogeneity of the effects on primary care utilization (section 6.2.1) and examining reform effects on the ACG weight and secondary care utilization (section 6.2.2). Finally, we compare the development in Östergötland to the corresponding developments in three other regions, that did not change their risk adjustment model at the same time (section 6.2.3).

6.1 Primary care utilization

6.1.1 Physician visits

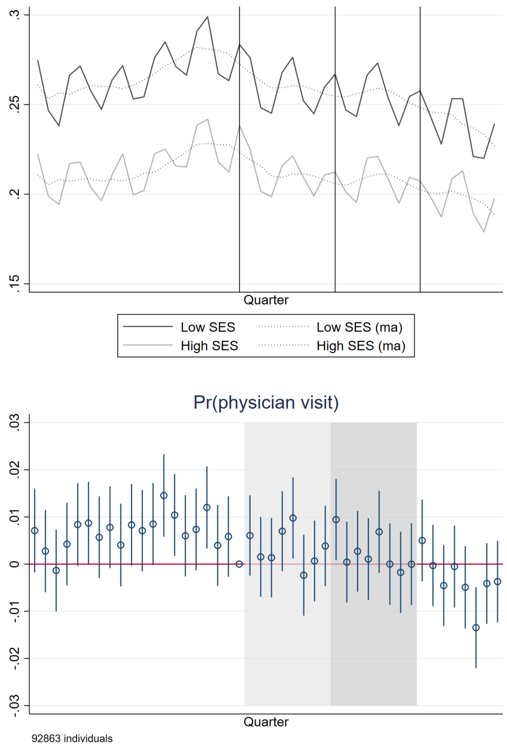
Figure 3 shows raw trends (top panel) and event-study estimates (bottom panel) for physician visits. The figures on the left show results for the probability of making at least one visit and the figures on the right show results for the number of physician visits. The upper figures display raw quarterly averages, together with moving averages calculated using a four-quarter window with two lags and one lead around each quarter. The vertical lines/shaded areas indicate the start of the announcement period (2012-2013), the phase-in period (2014-2015), and the period after the reform was fully implemented (2016-2017).

The moving averages show that the levels of both variables were stable over the period, except for a bump starting around the introduction of the patient choice system in the latter part of 2009. After 2015, the variables display a negative trend.

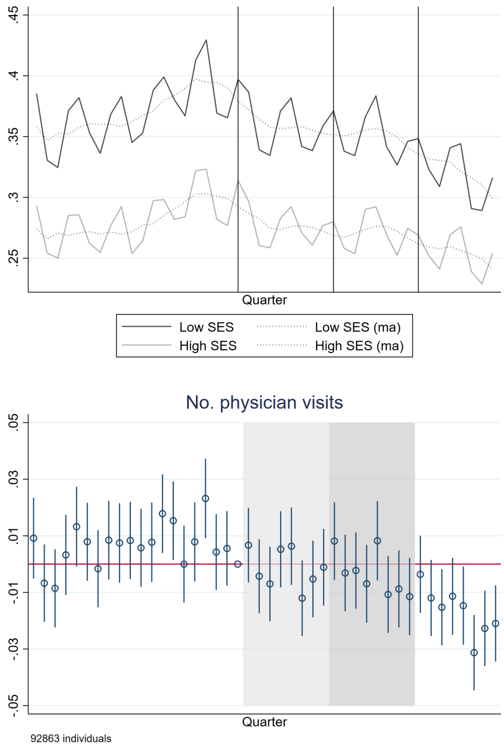
The utilization of primary care is generally higher in the low-SES group (black lines) than in the high-SES group (gray lines), which is expected given the lower health status in the low-SES group. Overall, the outcomes for the low- and high SES groups develop similarly over time. Importantly, the trends do not diverge in the way they would have done if the health status of low-SES individuals had deteriorated relatively faster during the study period (assuming worse health would imply increased utilization). By contrast, the most noticeable deviation between the two groups is that the number of visits *decreased more* in the low-SES group in the very last part of the study period. Despite the overall similarity, the event-study estimates from Eq. 1, shown in the lower part of Figure 3, indicate that the quarter estimates are generally more positive for the low-SES group during the pre-period, especially for the probability of a visit. However, as seen from the figure, this is solely driven by chosen reference quarter – if we had used any other pre-reform quarter as our reference period, almost all pre-period event-study estimates would have been very close to zero. In most years, the event-study estimates further display a seasonal pattern, with a larger SES difference in the first quarter of the year.[15] In particular, the event-study estimates for the first quarters are especially large in 2010 and 2011 (right before the announcement period). This pattern likely arises mechanically from the temporary increase in visits those years.

Nothing in these figures suggests that the payment reform had a positive impact on the access to physicians for low-SES individuals. From the announcement period onwards, the event-study estimates hover around zero. Although the two groups display seasonal differences, it is implausible that neutralizing shocks affecting the two groups differentially would occur in close to all post-announcement quarters.

[15] When we asked administrators in the region about the seasonal pattern, they speculated that it may follow mechanically from the tendency that physicians set the end date of sickness certificates to Jan 1, triggering visits due to renewals in the first quarter of the year. We lack data on the sickness absence among our study population, but the relatively worse health and large uptake of disability pension suggests that the low-SES group have higher sickness absence too.



(a) Probability of visit



(b) Number of visits

Figure 3: Physician visits by quarter 2007-2017

Note: Upper panel: The lines show raw quarterly averages by group (low SES/high SES) and the dots show moving averages over four quarters (two lags, 1 lead). *Lower panel:* Event-study estimates and 95% confidence intervals using the last quarter of 2011 as the reference quarter. The vertical lines (upper panel) and shaded areas (lower panel) indicate the announcement period (2012-2013) and the two first post reform years (2014-2015). The reform was fully phased in by 2016 (rightmost part of figures).

From the event-study estimates, we expect that formal DiD estimates would be zero or even negative. This is confirmed by the results shown in Table 3. The table displays DiDs from estimations of Eq. (2), separate for the three subperiods 2012-2013 (announcement period), 2014-2015 (phase-in period) and 2016-2017 (after the reform was fully phased in). Panel A displays the results for the probability of making at least one visit, and Panel B the results for the number of visits. In our preferred specification (column 1), the estimates are negative and statistically significant from the announcement period and onwards. The announcement and phase-in period estimates are not statistically different from each other, while the estimate for the last period is significantly more negative. The estimates for the probability of a visit correspond to a decrease of 1.4% (announcement), 2.0% (phase-in), and 4.6% (post period) of the mean in the high-SES group, and the estimates for the number of visits correspond to decreases of 2.8%, 3.5%, and 8.1% relative to the mean.

In columns 2-8, we vary the specification to examine the stability of the results. Column 2 shows that the inclusion of group-specific linear trends yields estimates that are even more negative than in the baseline specification. It is reassuring that the inclusion of a trend does not change our conclusions, given the presence of positive pre-reform event-study estimates in Figure 3. In this regard, we also note that a synthetic difference-in-differences model estimated on a 30% random subsample yields similar results (see Appendix D).

Column 3 of Table 3 shows that models not using the CEM weights yield smaller or even positive point estimates for the announcement and first post-reform periods. This is to be expected given the demographic differences between the high and low-SES groups. Nonetheless, even the positive estimates are precise enough to rule out meaningful increases in the low-SES group. From 2016, the estimates without weights are negative and statistically significant. Column 4 excludes the individual fixed effects, leaving the estimates virtually unchanged. Column 5 excludes individuals who left the sample during follow-up (due to migration or death). The estimates are smaller, implying that the negative estimates in the preferred specification partly reflect attrition. However, all estimates are negative, and the estimate for 2016-2017 is statistically significant for the probability of a visit (Panel A), as are all the estimates for the number of visits (Panel B).

Column 6 indicates that the estimates are slightly less negative when we use the predicted treatment status to compute a time-varying CNI measure; however, the differences to the preferred specification are small.[16] Column 7 shows results from a specification using a time-varying study population (see footnote 14). This is the only specification in which the estimates are consistently positive and statistically significant. Importantly, as we show in Online Appendix I, these estimates are entirely driven by a diverging pre-trend; when including group-specific trends, the estimates are very close to zero and statistically insignificant. Column 8 shows that clustering the standard errors at the PCC level has minor effects; the only difference is that the estimate for the announcement period loses significance in the model for the probability of a visit. Column 9 restricts the pre-period to the quarters after Östergötland introduced its patient choice reform (i.e., after 2010). As expected given the event-study estimates in Figure 3, the estimates are more negative in this specification, but they are still similar to the preferred estimates.

To conclude, across a range of specifications, we find no robust evidence that the payment reform increased the utilization of physician services for low-SES individuals relative to high-SES individuals – if anything, their utilization decreased.

[16]To predict CNI status in 2011, we estimate logit models for the potentially time-varying CNI components, i.e., indicators for being a *single parent, unemployed* or *above 25 and with no more than primary education.* Each model includes the first and the third lag of the dependent variable, interacted with a gender dummy, and a continuous birth year variable. We then use the estimates from these models to predict the probability of being a single parent/unemployed/having low education, and classify individuals as having a high probability if the predicted probability is at least 90 percent. We then create a predicted CNI variable as an indicator equal to one for individuals who had a high predicted probability for at least one of these components, or were born abroad (time-invariant characteristic). This approach assigns more individuals into the high-CNI group than the usual definition. Most of the mispredictions come from the very youngest birth cohorts; especially, the prediction approach assigns all women born after 1995 to the treatment group. This implies that these individuals are excluded from the sensitivity analysis using predicted CNI. A likely reason for the mispredictions among young people is that there is only one relevant time-varying CNI component for them: being a single parent. During our study period, they were too young to be counted as unemployed or low educated.

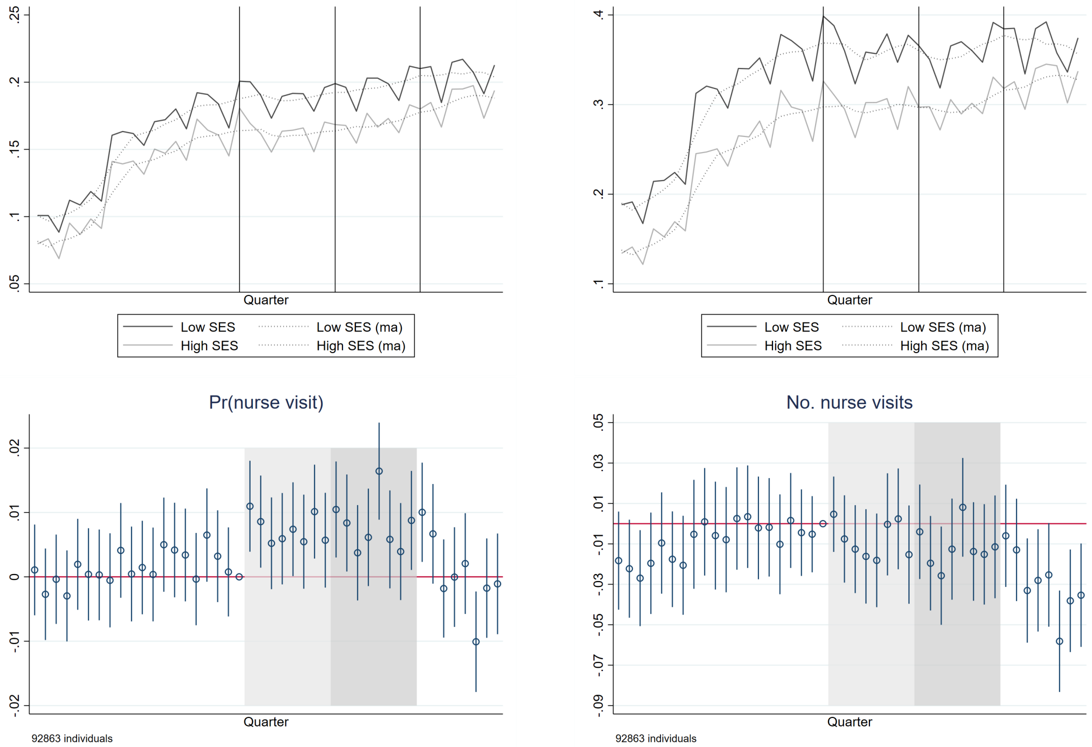
Table 3: Physician visits: Difference-in-differences estimates and robustness

	Panel A: Probability of visit							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Announcement	-0.00309*	-0.00457*	0.0000420	-0.00283	-0.00203	0.00652***	-0.00309	-0.00403*
	(0.00145)	(0.00226)	(0.00140)	(0.00150)	(0.00142)	(0.00136)	(0.00164)	(0.00164)
2014 reform	-0.00423**	-0.00656*	0.000756	-0.00202	-0.00307*	0.0107***	-0.00423*	-0.00518**
	(0.00156)	(0.00330)	(0.00150)	(0.00159)	(0.00152)	(0.00151)	(0.00173)	(0.00177)
2016 reform	-0.00989***	-0.0131**	-0.00345*	-0.00653***	-0.00854***	0.0100***	-0.00989***	-0.0108***
	(0.00160)	(0.00436)	(0.00154)	(0.00162)	(0.00156)	(0.00160)	(0.00239)	(0.00183)
N	4085972	4085972	4085972	3771196	3884012	4861192	4085972	2971616
No. clusters	92863	92863	92863	85709	88273	148180	44	92863
Mean dep.	0.215	0.215	0.215	0.219	0.217	0.220	0.215	0.214
Announcement=2014 DiD	0.483	0.321	0.648	0.631	0.517	0.00630	0.522	0.483
2014 DiD = 2016 DiD	0.000	0.001	0.006	0.006	0.000	0.661	0.006	0.000
Linear trend	No	Yes	No	No	No	No	No	No
CEM	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Attritioners	Included	Included	Included	Excluded	Included	Included	Included	Included
Treatment def.	Fixed	Fixed	Fixed	Fixed	Predicted	Yearly	Fixed	Fixed
Pop. def.	Cohort	Cohort	Cohort	Cohort	Cohort	Yearly	Cohort	Cohort
Cluster	Ind	Ind	Ind	Ind	Ind	Ind	PCC	Ind
Min year	2007	2007	2007	2007	2007	2007	2007	2010
	Panel B: Number of visits							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Announcement	-0.00785**	-0.0133***	-0.00408	-0.00734**	-0.00591*	0.00737**	-0.00785**	-0.0107***
	(0.00248)	(0.00382)	(0.00238)	(0.00256)	(0.00241)	(0.00227)	(0.00281)	(0.00276)
2014 reform	-0.00979***	-0.0184**	-0.00384	-0.00593*	-0.00749**	0.0130***	-0.00979**	-0.0126***
	(0.00268)	(0.00567)	(0.00258)	(0.00275)	(0.00259)	(0.00256)	(0.00326)	(0.00304)
2016 reform	-0.0229***	-0.0346***	-0.0162***	-0.0175***	-0.0200***	0.00667*	-0.0229***	-0.0257***
	(0.00269)	(0.00751)	(0.00259)	(0.00275)	(0.00261)	(0.00264)	(0.00424)	(0.00310)
N	4085972	4085972	4085972	3771196	3884012	4861192	4085972	2971616
No. clusters	92863	92863	92863	85709	88273	148180	44	92863
Mean dep.	0.283	0.283	0.283	0.289	0.287	0.289	0.283	0.282
Announcement=2014 DiD	0.472	0.137	0.928	0.613	0.548	0.0280	0.541	0.472
2014 DiD = 2016 DiD	0.000	0.000	0.000	0.000	0.000	0.013	0.000	0.000
Linear trend	No	Yes	No	No	No	No	No	No
CEM	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Attritioners	Included	Included	Included	Excluded	Included	Included	Included	Included
Treatment def.	Fixed	Fixed	Fixed	Fixed	Predicted	Yearly	Fixed	Fixed
Pop. def.	Cohort	Cohort	Cohort	Cohort	Cohort	Yearly	Cohort	Cohort
Cluster	Ind	Ind	Ind	Ind	Ind	Ind	PCC	Ind
Min year	2007	2007	2007	2007	2007	2007	2007	2010

Note: In each model, the dependent variable is the either the an indicator for at least one visit (Panel A) or the number of physician visits (Panel B) at a primary care center in a quarter. The table shows three average DiDs, each computed over a number of quarterly differences-in-differences (DiD) estimates from linear regression models. The estimates contrast individuals with Care Need Index (CNI)>0 to individuals with CNI=0; Announcement shows the average of the quarterly DiD estimates in the announcement period (Q2 2012 – Q4 2013), 2014 reform shows the average of the quarterly DiD estimates after the reform (2014-2015), and 2016 reform shows the average of the quarterly DiD estimates after the price increase in 2016 (2016-2017). Announcement=2014 DiD shows p-value of test of equality of the DiDs for 2012-13 and 2014-15 periods. 2014 DiD = 2016 DiD shows p-value of test for equality of DiDs in 2014-15 and 2016-17. All models include quarter fixed effects. Individual's CNI status is measured on Dec 31 2011 unless stated otherwise (see row Treatment def). Predicted = use predicted CNI for 2011 instead of actual CNI. CEM weights balance the sample in terms of birth year and gender. Attritioners are defined as individuals moving out from the region or dying in 2013-2017. The study population is a cohort who were registered at a PCC in the region on Jan 1 2013 and lived in the region throughout 2007-2011; the exception is the column for which Pop.Def is indicated as Yearly. In that specification, the study population includes individuals registered at a PCC in the region on January 1 year t, for t=2010-2017. Standard errors are clustered by individual in all models except in column 8, in which they are clustered by the individual's PCC at baseline. * p<0.05, ** p<0.01, *** p<0.001.

6.1.2 Nurse visits

Nurses play a prominent role in Swedish primary care. We therefore also report trends and event-study estimates for the probability and number of *nurse* visits, although with the disclaimer that the contemporaneous reorganization of home care may have had spillover effects also on our study sample. Figure 4 shows the development of nurse visits over the study period. The upper part of the figure shows large increases in the levels of the two outcomes in very first part of the study period, and a slower but still increasing trend thereafter. The event-study estimates in the lower part of the figure do not indicate differential pre-trends for the probability of seeing a nurse, and the pre-trends for the number of visits move closely together after the structural break in the early study period. We note that the probability of seeing a nurse increased more for the low-SES group in the announcement period and in 2014-2015, but the average number of visits did not. In 2016-2017, the high-SES group approached the low-SES group both on the extensive (probability of visit) and intensive (number of visits) margin.



(a) Probability of visit

(b) Number of visits

Figure 4: Nurse visits by quarter 2007-2017

Note: Upper panel: The lines show raw quarterly averages by group (low SES/high SES) and the dots show moving averages over four quarters (two lags, 1 lead). *Lower panel:* Event-study estimates and 95% confidence intervals using the last quarter of 2011 as the reference quarter. The vertical lines (upper panel) and shaded areas (lower panel) indicate the announcement period (2012-2013) and the two first post reform years (2014-2015). The reform was fully phased in by 2016 (rightmost part of figures).

Table 4: Nurse visits: Difference-in-differences estimates

	Probability of visit	Number of visits
	(1)	(2)
Announcement	0.00611***	0.000573
	(0.00143)	(0.00698)
2014 reform	0.00665***	-0.00332
	(0.00159)	(0.00713)
2016 reform	-0.000807	-0.0212**
	(0.00172)	(0.00777)
N	4085972	4085972
Individuals	92863	92863
Mean dep.	0.145	0.257
Announcement=2014 DiD	0.728	0.581
2014 DiD = 2016 DiD	0.000	0.001

Note: The dependent variable is either an indicator for at least one visit or the quarterly number of visits with a nurse at a primary care center. The estimates in the table are average DiDs, computed over a number of quarterly differences-in-differences (DiD) estimates from linear regression models. The estimates contrast individuals with Care Need Index (CNI)>0 to individuals with CNI=0; Announcement shows the average of the quarterly DiD estimates in the announcement period (Q2 2012 – Q4 2013), 2014 reform shows the average of the quarterly DiD estimates after the reform (2014-2015), and 2016 reform shows the average of the quarterly DiD estimates after the price increase in 2016 (2016-2017). Announcement=2014 DiD shows p-value of test of equality of the DiDs for 2012-13 and 2014-15 periods. 2014 DiD = 2016 DiD shows p-value of test for equality of DiDs in 2014-15 and 2016-17. All models include quarter fixed effects. Individual's CNI status is measured on Dec 31 2011. CEM weights balance the sample in terms of birth year and gender. Attritioners are defined as individuals moving out from the region or dying in 2013-2017. The study population is a cohort born 1953-2001 who were registered at a PCC in the region on Jan 1 2013 and lived in the region throughout 2007-2011 and had a chronic condition status registered in the pre-period. Standard errors are clustered by individual. * p<0.05, ** p<0.01, *** p<0.001.

Table 4 shows the DiD estimates for nurse visits. Looking first at the probability of a visit (column 1), the DiDs are positive and statistically significant in the announcement and phase-in periods (amounting to 4.2 and 4.6% of the control group mean), but negative, though small and statistically insignificant, in 2016-2017. The DiD estimates for the number of nurse visits (column 2) are small and statistically insignificant in the announcement and phase-in periods, and negative and statistically significant in 2016-2017 (8.2% of the control group mean). In Appendix E, we report robustness tests similar to the ones used for physician visits. The results are generally stable, except that the estimate for the probability of a visit loses significance when we include a linear trend or cluster the standard errors at the PCC level. By contrast, a synthetic DiD model of the probability reported in Appendix D suggests that the positive effect may have lasted throughout the entire post-period. While the synthetic DiD approach yields positive effects also for the number of visits, this result appears to be driven by regression to the mean.

6.1.3 Summary primary care utilization

Across a range of specifications, we fail to find any robust evidence suggesting that primary care providers responded to the payment reform by offering low-SES individuals increased access to physicians. We do find some evidence that the likelihood of seeing a nurse increased more for low-SES individuals, at least temporarily, but the result is not robust to including a linear trend. The results underscore the point made in the theoretical framework: an increase in the capitation for a given patient or patient group need not imply an increase in services provided for that same patient or group.

For both physician and nurse visits, the DiD estimates are more negative after the reform was fully phased in (2016-2017). The event study graphs show that this is mostly driven by 2017, the year when the region increased the patient fees. It is plausible that the negative estimates for 2017 to some extent reflect that low-SES individuals are more sensitive to the fee level; consequently, we do not take the more negative effect for 2016-2017 to indicate a dose-response relationship. Importantly, the conclusion that the reform did not increase the access to primary care of low-SES patients does not depend on whether we include 2017 in the sample or not.

6.2 Extensions

To further understand the lack of reform effects, we present three extensions of the analysis below. Section 6.2.1 examines if the main results hide heterogeneity across patient and provider characteristics that theoretically may produce opposing reactions to the reform or otherwise explain why a behavioral response may be small. Section 6.2.2 explores downstream effects of the reform on the individual's ACG weight and secondary care utilization. Finally, in section 6.2.3 we make a counterfactual analysis by comparing our results to similar estimates from three other Swedish regions, which did not change their risk adjustment models in the same year.

6.2.1 Heterogeneity across patient and provider characteristics

There are several aspects of our study setting that, theoretically, may explain why we find little effect of the reform on the SES differential. A first potential explanation for the lack of increase in primary care utilization stems from the SES adjustment used in the old risk-adjustment model (see section 3.1). For low-SES individuals living in the very poorest areas, the reform did not increase the capitation. To explore if such heterogeneity in the reform impact on the capitation explains our main results, we exclude the patients enrolled at the two PCCs which benefited the most from the *old* SES-adjustment from the estimation sample. The DiD estimates do not suggest that the main results are explained by opposing effects on individuals living in the poorest areas and individuals living elsewhere (Panel A, Table F.1). For physician visits, the estimates are very similar. For nurse visits, the positive estimates on the probability of a visit in the 2012-2015 are somewhat smaller than in the main specification.

A second, and related, potential explanation is that the introduction of ACG adjustment reduced the capitation of low-SES individuals in good health (i.e., people with low ACG). As seen from Figure 2, it is relatively common to have a low ACG weight in a given year, even in our study population of people with a chronic condition. To see if the reform had a more positive impact on the utilization for individuals with poor health status, we estimate our main models on a restricted sample that only includes observations whose ACG weight ≥1. To closely approximate the incentives in a given quarter, we use the contemporary ACG weights to make the restriction (instead of, e.g., the value in 2011). With this restriction, almost everyone in the estimation sample would receive a higher capitation due to the reform. Nonetheless, we obtain similar results as in the main specification for the physician visit outcomes (Panel B, Table F.1). The positive estimates on the probability of seeing a nurse become smaller and statistically insignificant when we remove observations with ACG < 1. Thus, the individuals with the poorest health did not get greater access to nurses. Possibly, the increased access to nurses was targeted to relatively healthy individuals; alternatively, the result could indicate reverse causality: nurses tend to register fewer diagnoses, thus individuals may receive a lower ACG weight in years when they are especially likely to see a nurse instead of a physician.[17] To conclude, the estimates in our main specifications are *not* driven by low-SES individuals with relatively good health status, whose capitation might have decreased after the reform.

A third possible explanation for the lack of effects in the main analysis is that providers may not be able to distinguish between patients with low or high SES. That is, even if they would want to provide more care to low-SES patients following the introduction of risk adjustment, they may struggle to identify the relevant patient group. While this may be a plausible explanation for SES aspects such as educational attainment and single-parenthood, it is less plausible when it comes to non-EU immigrant background, which should be a visible characteristic in the Swedish context. Thus, if this explanation

[17] Notably, having an ACG weight below 1 is a transitory phenomenon in our study population: 89,483 of the individuals in our total study population (92,863 individuals) remain in the estimation sample, though they appear in much fewer quarters. An ACG weight equal to 1 implies having exactly average health care costs. Given the skewness of health care costs, it would rather be surprising if people had an ACG weight above 1 most years.

is valid, we would expect a clear heterogeneity in terms of the different dimensions composing the CNI (immigrant background, single parent, short education, or unemployed). Yet, event study models estimated by subgroup does not indicate such heterogeneity (Online Appendix G).

A fourth possible explanation for the lack of positive effect on utilization may be that many PCCs are publicly owned, and as such may have softer budget constraints or just a weaker incentive structure than private PCCs. That is, the null results may reflect that the public PCCs did not respond to the reform. When examining heterogeneity by ownership type, we do however find a similar development for both private and public PCCs (Online Appendix H). In fact, the DiD estimates indicate that the effects are significantly more negative for private PCCs. The estimates for private PCCs indicate that increasing the care provided to low-SES patients was not the profit-maximizing response to the reform.

A fifth explanation relates to the fact that risk adjustment entails a redistribution of existing funds, meaning that the net effect on the funds available for a given PCC may be small. In such cases, the net effect of receiving lower capitation for some individuals and higher capitation for others may cancel out. In other words, PCCs who were expected to gain or lose money due to the redistribution may respond more strongly to the reform. Furthermore, if PCCs who would gain or lose from the reform behave symmetrically, the lack of improvements in access may be driven by opposing effects. To explore this possibility, we use data on the expected budget change between 2013 and 2014 for each PCC. The negative event-study patterns are similar for PCCs whose budgets were expected to increase, decrease, or remain approximately stable (Online Appendix H), suggesting that opposing effects on the capitation within a given PCC do not explain the lack of effects. Notably though, the DiD model indicates a significantly more negative estimate for PCCs whose budget was projected to increase due to the reform. The pattern that PCCs gaining more funds for low-SES patients respond by providing less care is thus stable.

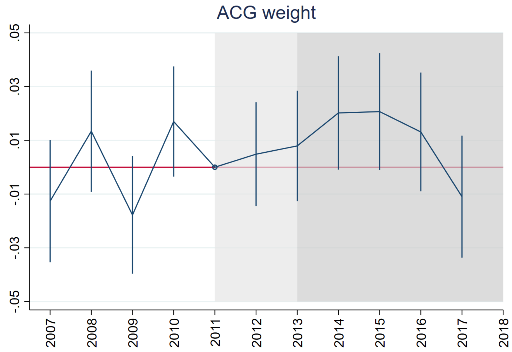
To summarize, the heterogeneity analyses do not suggest that the small impact on primary care utilization is driven by opposing effects on the capitation or by providers struggling to identify low-SES patients. While the null effect is consistent with a lack of profit motive for public providers, the results do not indicate that the private providers respond drastically different – and certainly not in a way that would increase the utilization.

6.2.2 Additional outcomes: ACG weight and secondary care

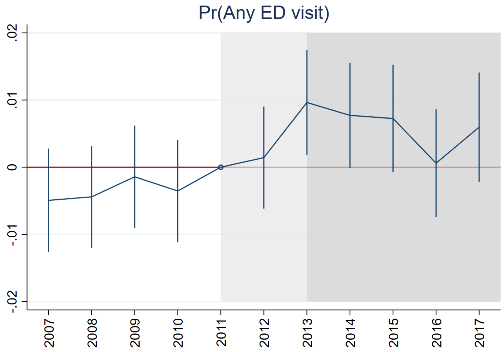
Although the reform did not have any great impact on the probability or number of primary care visits, it is possible that it affected the quality of care. It is therefore interesting to examine other outcomes that may be linked to patients' health and downstream use of other types of care.

We first consider the individual ACG weight as an outcome variable. Reform-induced changes to the ACG weight reflects two, not mutually exclusive, phenomena: changes in health and changes in the number and type of providers visited (physicians and nurses, as well as primary and secondary providers, have different diagnosis registration patterns). Figure 5(a) shows event-study estimates for the ACG weight, and column 1 of Table 5 shows the corresponding DiD estimates. The event-study estimates are positive in the announcement and phase-in period, become smaller in 2016, and turn negative in 2017, although none of the estimates are statistically significant. The DiD specification indicates that there is a significant positive effect in 2014-2015. The estimates are small, and thus do not indicate substantial effects of the reform.

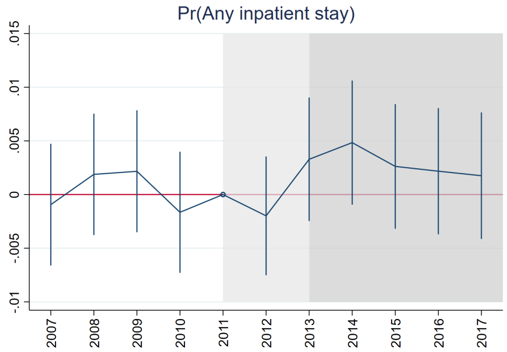
The reform might have affected individuals' propensity to seek care outside the primary care setting. Figure 5(b) shows estimates on the probability of having visited a hospital emergency department during a year. Column 2 of Table 5 shows the corresponding DiD estimates. The event study estimates are positive all post-announcement years, and significant in 2013-2014. The DiD estimates are significant in all post periods. Figures 5(c)-(d) shows event-study estimates of the probability of a hospitalization and the number of hospital days. The corresponding DiD estimates are shown in Column 3 and 4 of Table 5. The estimates indicate positive but small and statistically insignificant effects on these outcomes. Figure 5(e) shows event-study estimates of the probability of an avoidable hospitalization in a given year. Column 5 of Table 5 shows the corresponding DiD estimates. The sign of the event-study estimates varies between periods. In the DiD specifications, all estimates are positive but they are close to zero and not statistically significant.



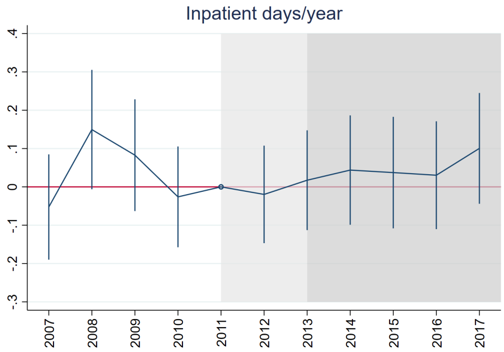
(a) ACG



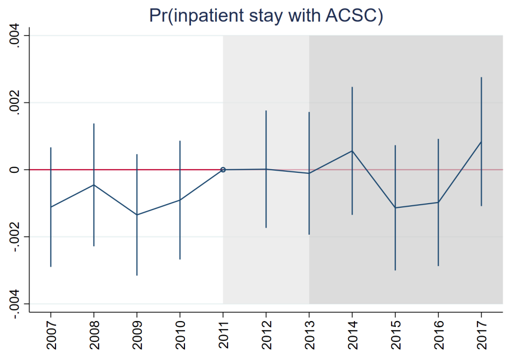
(b) ED



(c) Any hospitalization



(d) Hospital days



(e) ACSC hospitalization

Figure 5: Other outcomes; annual data 2007-2017

Note: Event-study estimates and 95% confidence intervals for other outcome variables. ACG weight = risk score from the morbidity risk-adjustment model. Any ED visit = indicator for having at least one visit at an emergency department at hospital in Östergötland. Pr(Any inpatient stay) = indicator for having at least one inpatient stay at any hospital in Sweden. Inpatient days/year = number of days in hospital during the year (any Swedish hospital). Any ACSC = indicator for having at least 1 hospitalization with an ambulatory care sensitive condition diagnosis (any Swedish hospital).

Table 5: Other outcomes (difference-in-differences estimates)

	(1) ACG weight	(2) Any ED visit	(3) Any inpatient stay	(4) Inpatient days	(5) Any ACSC
Announcement	0.00639	0.00837***	0.000342	-0.0320	0.000719
	(0.00689)	(0.00241)	(0.00176)	(0.0459)	(0.000555)
2014 reform	0.0205**	0.0103***	0.00344	0.00986	0.000482
	(0.00751)	(0.00251)	(0.00185)	(0.0495)	(0.000606)
2016 reform	0.00114	0.00613*	0.00167	0.0345	0.000691
	(0.00801)	(0.00260)	(0.00191)	(0.0519)	(0.000616)
Constant	1.294***	0.187***	0.0842***	0.782***	0.00730***
	(0.00103)	(0.000319)	(0.000239)	(0.00656)	(0.0000762)
N	1009057	1003660	1003631	1003631	1003929
Individuals	92863	92863	92863	92863	92863
Announcement=2014 DiD	0.0607	0.490	0.132	0.428	0.723
2014 DiD = 2016 DiD	0.0116	0.148	0.392	0.651	0.761

Note: ACG weight = risk score from morbidity risk adjustment model. Any ED visit = Pr(at least one visit at an emergency department (hospitals in Östergötland). Any inpatient = Pr(at least one inpatient stay (any Swedish hospital). Inpatient days = no. inpatient days (any Swedish hospital). Any ACSC = indicator for having at least 1 hospitalization with an ACSC diagnosis (any Swedish hospital). The estimates contrast individuals with Care Need Index (CNI)>0 to individuals with CNI=0; Announcement shows the DiD estimate for the announcement period (2012 – 2013), 2014 reform shows the DiD estimate after the reform (2014-2015), and 2016 reform shows the DiD estimate after the price increase in 2016 (2016-2017). Announcement=2014 DiD shows p-value of test of equality of the DiDs for 2012-13 and 2014-15 periods. 2014 DiD = 2016 DiD shows p-value of test for equality of DiDs in 2014-15 and 2016-17. All models include individual and year fixed effects, and use CEM weights to balance the sample in terms of birth year and gender. The constant shows the mean of the dependent variable in the low-CNI group. Standard errors are clustered by individual. * p<0.05, ** p<0.01, *** p<0.001.

The small and statistically insignificant effects on hospitalizations suggest that the reform did not affect the likelihood of experiencing an adverse health event. Given the lack of effects on hospitalizations and the negative estimate on primary care physician visits, it further seems unreasonable to interpret the positive estimates for ACG and ED visits as reflecting deteriorations in health. A more plausible interpretation is that low-SES patients responded to decreased access to primary care physicians by turning to the ED. Differences in diagnosis registration patterns may then explain the positive estimates on the ACG weight. Note that we find the largest effect on the ACG weight in 2014-2015, i.e., the years with the largest estimates for the probability of visiting the ED.

6.2.3 Comparison with other regions

As mentioned in section 5.1, the major threat to our identification strategy is that the health of the low-SES group may have deteriorated more over the study period for reasons that are unrelated to the reform. To examine the scope for this type of bias, we use data from three other Swedish regions to estimate corresponding event study models as previously shown for Östergötland. If health deteriorates faster for low-SES than for high-SES individuals, the SES difference should be growing throughout the period, with the low-SES group either increasing more or decreasing less than the high-SES group over time. Empirically, we do not find such a pattern for neither physician nor nurse visits in primary care (Figure J.1), the hospitalization outcomes (Figure J.2), or the ACG (Figure J.3) in the comparison regions.[18] These results

[18] By contrast, we note that for individuals with no pre-existing chronic condition, the ACG weight increased considerably more for low-SES individuals in the post reform years. That is, the risk of developing or having undetected illnesses is larger in the low-SES group. This pattern suggests that the approach of splitting the sample by the pre-existence of a chronic condition introduces mean reversion in the subsample *without* such a condition (see e.g., Daw and Hatfield, 2018; Illenberger et al., 2020, for how matching may induce mean reversion) and therefore that a DiD strategy is inadequate to study that group, because it would mechanically produce positive estimates. By contrast, in our study population of individuals with a chronic condition,

for other regions show that faster health deteriorations in the low-SES group are not inevitable – at least not over periods of the same length as our study period, and for individuals in the studied age groups.

Although there is no evidence of *growing* divergence, we note that there are small one-off positive shifts of some of the variables (physician visits, nurse visits, and the ACG weight) in the comparison regions. Ironically, the low-SES groups in these regions thus saw a more favorable development than the low-SES group in our study region. This informal triple-DiD provides further suggestive evidence that the risk-adjustment reform in Östergötland did not improve the primary care access for low-SES individuals, although this evidence should be taken with a grain of salt due to substantial changes to the market structure in 2008-2009 in the three comparison regions and a payment system reform in 2016 in one region.

7 Discussion and concluding remarks

We study how the introduction of morbidity risk adjustment and a more sensitive SES-based adjustment of the capitation to primary care providers affects socioeconomic differences in care utilization. Focusing on a core group of patients for primary care providers – individuals with chronic conditions – we find no evidence of increased access to primary care physicians for individuals with low SES relative to individuals with high SES. If anything, our estimates suggest the opposite. While there are some signs of a temporary increase in the likelihood of visiting a primary care nurse, this result is not robust to specification changes, and the result is completely absent for the patients with the poorest health. Thus, for the patients with the lowest health status, we find no improved access to either primary care physicians or nurses after the reform. Furthermore, the temporary substitution of physician and nurse visits of more healthy low-SES patients can hardly be interpreted as improved access.

The conclusion that access did not improve is robust to a range of specification changes. Furthermore, the main results do not hide a mix of positive and negative effects on subgroups whose capitation was more and less positively affected by the reform, or on groups whose SES would be more or less observable. We find more negative estimates for treated individuals enrolled at private PCCs or at PCCs whose total revenues (budget) was projected to increase after the reform, although the overall pattern is quite similar across PCC types.

The point estimates for physician visits become even more negative after the reform was fully phased in. Notably, the negative estimates in the very last year (2017) may partly reflect that low- and high-SES individuals reacted differently to a concurrent increase in the patient fee. If so, it is a cruel irony that the policy intended to address socioeconomic inequalities in health (risk adjustment) failed to do so, while another policy designed for other purposes (the fee increase) might have aggravated the inequalities.

We also study outcomes from the secondary care sector. We find a positive effect on the probability of visiting a hospital emergency department. As the ED may function as a substitute for primary care,

greater health deterioration for the low-SES group seem less plausible (as also shown in this section).

this result is consistent with our conclusion that the reform did not improve the access to primary care. We do not find any effect on hospitalizations, which is the best proxy for health status we can obtain in our administrative data. These results do not suggest that the reform had a meaningful effect on socioeconomic health inequalities.

Our results are unlikely to be explained by poor fit of the new risk-adjustment model. The increase of the average capitation in the low-SES group was large, and substantially larger than the increase in the high-SES group. Furthermore, there is little in this context suggesting that the lack of response reflects that primary care providers fail to observe patients' SES. While a patient's educational attainment or single-parent status may not be apparent, it should be possible to observe if the patient has immigrated from a country outside the European Union in our setting. Yet, our results are similar regardless of if we consider education, single-parenthood, unemployment or immigrant status as the definition of SES. In the next section, we discuss the plausibility of mechanisms that we could not test directly with our data.

7.1 Potential mechanisms

The theoretical framework in Section 2 offers a range of explanations why the risk adjustment reform had limited effects on the socioeconomic differences in care utilization and health. As we discussed in section 3.4, some of these seem implausible in our setting. *i)* With most savings from prevention falling outside the primary care sector, the results are unlikely to reflect diluted incentives for prevention. *ii)* Although GPs and nurses are salaried, several previous studies document that PCCs react to financial incentives, suggesting that our results do not reflect weak incentives. Further, whilst the heterogeneity analysis in section 6.2.1 shows that the estimate is more negative for private providers, the pattern for public providers is similar, not opposite. Thus, a relatively weak profit motive of public providers is, at best, a partial explanation of our results. *iii)* The reform plausibly signalled to providers that they should provide more care to low-SES patients. Thus, it is unlikely that the null effects are explained by the reform confirming providers' prior beliefs about the priority of different patients.

The fundamental reason for the lack of effects is that a prospective payment such as capitation gives providers autonomy over the allocation of funds. Their choices, in turn, depends on their objectives. The semi-altruistic providers may have retained the increased capitation as profits, or channelled the funds towards other patients with a higher responsiveness to care quality (competition mechanism) or even higher care needs (altruism mechanism) (Barham and Milliken, 2015). The altruism explanation is consistent with the overall decrease in primary care use among both low- and high SES patients in our study population. It is also consistent with the larger negative effect on providers whose budget increased relatively much after the reform.

7.2 Limitations

The main threat to the parallel trends assumption needed to identify the effects on socioeconomic differences in care utilization is that low-SES individuals' health may have deteriorated more quickly than high-SES individuals' health. We restricted the sample to individuals with chronic conditions to mitigate this risk and found little evidence of a differential development. A potential confounding factor is the change in patient fees implemented in 2017, which may have affected low- and high-SES individuals differently. It is difficult to separate the effects of the new risk-adjustment model and this change; however, the conclusion that the new risk-adjustment model did not improve access holds also if we exclude 2017 from our study period.

It deserves to be emphasized that increasing the number of visits is not the only way in which the PCCs might have reacted to the risk adjustment. Low-SES patients may have benefited in other ways that we could not observe, such as extended consultation time. Further, we cannot rule out that providers engaged in outreach activities not registered in the administrative data, although anecdotal evidence from another Swedish region with a similar risk-adjustment model suggests that such activities are rare (Anell, 2016a,b). In any case, we note that any such efforts evidently did not make enough of a difference to reduce the incidence of adverse health events (as indicated by hospitalizations). It therefore seems most appropriate to conclude that the payment reform did not lead to meaningful improvements for the low-SES individuals. Further, the decreasing visit trends for both the low- and high SES groups indicate that the reform did not help patients with chronic conditions, irrespective of their SES. It is possible that other groups benefited, but such effects are outside the scope of this analysis of the impact on socioeconomic inequalities.[19]

As we cannot measure all provider activities, we cannot determine whether the null effects are driven by a lack of supply-side response or a lack of demand. However, our results do show that any outreach activities that may have taken place evidently were not powerful enough to attract low-SES patients with high need according to providers. Thus, we note that even a large redistribution of funds may not be sufficient to reduce inequalities.

7.3 Concluding remarks

Despite the widespread use and the theoretically ambiguous consequences of risk-adjusted payment in health care, there are very few empirical studies of the effects of risk-adjustment models. We show that the introduction of morbidity risk adjustment and a more sensitive SES-based adjustment of the capitation to primary care providers did not affect socioeconomic differences in care utilization in a

[19]To conduct an evaluation of the total effects of the reform, we would have needed primary care data from a large number of comparison regions with stable primary care institutions over the study period. In Sweden, there is no national register of primary care akin to the national patient register for secondary care. Even if we would have approached all 21 Swedish regions to collect data for such an evaluation, considerable empirical challenges would remain due to the relatively small number of regions (clusters), which all implemented similar risk-adjustment models during a relatively short period of time. Our primary interest was the impact on socioeconomic inequalities, which allow for a within-region comparison.

sample of 6-64 year-olds with chronic conditions in a Swedish region.

Our results are consistent with a number of mechanisms. We believe that the most plausible explanations of the null effects are that low-SES individuals have relatively low responsiveness to quality, and that altruistic providers assign higher priority to other groups such as young children and elderly patients.

A message to policymakers is that risk adjustment of the capitation to primary care providers may not be an effective method to reduce socioeconomic inequalities in health. Coupled with the results from a related literature showing that physicians are less likely to undertreat high-need patients when their payment is based on the service volume than when payment is fixed (e.g., Hennig-Schmidt et al., 2011; Brosig-Koch et al., 2017; Cadena and Smith, 2022; Skovsgaard et al., 2023), our results suggest that it is preferable to link the payment more closely to the care actually provided to patients if the goal is to reduce inequalities. To the extent that payers use risk adjustment to signal the priority of different patient groups, it may be more effective to use other communication channels such as regular audit and feedback systems. The study also suggests the need for interventions targeting the demand side, such as reduced copayments and outreach activities to increase low-SES patients' responsiveness to quality and demand for primary care (Elinder et al., 2023; Sabety et al., 2023).

To further the literature, future studies should aim for research designs that can separate between competing mechanisms, and use more comprehensive measures of care utilization and quality. As the devil resides in the details when it comes to payment schemes, evidence from other institutional contexts is clearly warranted.

References

Abadie, A., 2005. Semiparametric difference-in-differences estimators. Review of Economic Studies 72, 1–19.

Abadie, A., Athey, S., Imbens, G.W., Wooldridge, J., 2023. When should you adjust standard errors for clustering? Quarterly Journal of Economics 138, 1–35.

Agerholm, J., Bruce, D., Ponce De Leon, A., Burström, B., 2013. Socioeconomic differences in healthcare utilization, with and without adjustment for need: An example from Stockholm, Sweden. Scandinavian Journal of Public Health 41, 318–325.

Aldstedt, K., 2012. Inriktningsbeslut avseende vårdval primärvård. Landstinget i Östergötland 2012-820.

Aldstedt, K., Melin, L.A., Paananen, C., Strömstedt, M., Löfsved, A.K.W., Zingmark, A.M., Öberg, S., 2015. Ersättningssystemet inom primärvården. HSN 2015-399.

Anell, A., 2015. The public-private pendulum – Patient choice and equity in sweden. New England Journal of Medicine 372, 1–4.

Anell, A., 2016a. Delrapport 2a: Bidrar CNI-ersättning till BVC till ett förändrat vårdutbud som skapar förutsättningar för en mer jämlik hälsa? mimeo.

Anell, A., 2016b. Översyn av primärvårdens utveckling efter införandet av Hälsoval Skåne. KEFU-Skånes skriftserie 2016:4.

Anell, A., Dackehag, M., Dietrichson, J., 2018. Does risk-adjusted payment influence primary care providers' decision on where to set up practices? BMC Health Services Research 18, 179.

Anell, A., Dackehag, M., Ellegård, L.M., 2021a. Weak association between socioeconomic Care Need Index and primary care visits per registered patient in three Swedish regions. Scandinavian Journal of Primary Health Care 39, 288–295.

Anell, A., Dietrichson, J., Ellegård, L.M., Kjellsson, G., 2021b. Information, switching costs, and consumer choice: Evidence from two randomized field experiments in Swedish primary health care. Journal of Public Economics 196, 104390.

Angerer, S., Waibel, C., Stummer, H., 2019. Discrimination in health care: a field experiment on the impact of patients' socioeconomic status on access to care. American Journal of Health Economics 5, 407–427.

Arkhangelsky, D., Athey, S., Hirshberg, D.A., Imbens, G.W., Wager, S., 2021. Synthetic difference-in-differences. American Economic Review 111, 4088–4118.

Baicker, K., Mullainathan, S., Schwartzstein, J., 2015. Behavioral hazard in health insurance. The Quarterly Journal of Economics 130, 1623–1667.

Bailey, M.J., Goodman-Bacon, A., 2015. The war on poverty's experiment in public medicine: Community health centers and the mortality of older americans. American Economic Review 105, 1067–1104.

Barham, V., Milliken, O., 2015. Payment mechanisms and the composition of physician practices: Balancing cost-containment, access, and quality of care: Payment mechanism and composition of physician practices. Health Economics 24, 895–906.

Barros, P.P., 2003. Cream-skimming, incentives for efficiency and payment system. Journal of Health Economics 22, 419–443.

Bilinski, A., Hatfield, L.A., 2018. Nothing to see here? Non-inferiority approaches to parallel trends and other model assumptions. arXiv preprint arXiv:1805.03273 .

Biørn, E., Godager, G., 2010. Does quality influence choice of general practitioner? An analysis of matched doctor–patient panel data. Economic Modelling 27, 842–853.

Brekke, K.R., Holmås, T.H., Monstad, K., Straume, O.R., 2019. How does the type of remuneration affect physician behavior?: Fixed salary versus fee-for-service. American Journal of Health Economics 6, 104–138.

Brosig-Koch, J., Hennig-Schmidt, H., Kairies-Schwarz, N., Wiesen, D., 2017. The effects of introducing mixed payment systems for physicians: Experimental evidence. Health Economics 26, 243–262.

Brown, J., Duggan, M., Kuziemko, I., Woolston, W., 2014. How does risk selection respond to risk adjustment? New evidence from the Medicare Advantage program. American Economic Review 104, 3335–64.

Cabral, M., Carey, C., Miller, S., 2021. The impact of provider payments on health care utilization of low-income individuals: Evidence from Medicare and Medicaid. Working Paper 29471. National Bureau of Economic Research.

Cabral, M., Geruso, M., Mahoney, N., 2018. Do larger health insurance subsidies benefit patients or producers? Evidence from Medicare Advantage. American Economic Review 108, 2048–87.

Cadena, B.C., Smith, A.C., 2022. Performance pay, productivity, and strategic opt-out: Evidence from a community health center. Journal of Public Economics 206, 104580.

Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A., Cutler, D., 2016. The association between income and life expectancy in the united states, 2001-2014. JAMA 315, 1750–1766.

Cookson, R., Propper, C., Asaria, M., Raine, R., 2016. Socio-economic inequalities in health care in England. Fiscal Studies 37, 371–403.

Dackehag, M., Ellegård, L.M., 2019. Competition, capitation, and coding: Do public primary care providers respond to increased competition? CESifo Economic Studies 65, 402–423.

Daw, J.R., Hatfield, L.A., 2018. Matching and regression to the mean in difference-in-differences analysis. Health Services Research 53, 4138–4156.

Devaux, M., 2015. Income-related inequalities and inequities in health care services utilisation in 18 selected OECD countries. European Journal of Health Economics 16, 21–33.

Devlin, R.A., Sarma, S., 2008. Do physician remuneration schemes matter? The case of Canadian family physicians. Journal of Health Economics 27, 1168–1181.

Dietrichson, J., Ellegård, L.M., Kjellsson, G., 2020. Patient choice, entry, and the quality of primary care: Evidence from Swedish reforms. Health Economics 29, 716–739.

Doorslaer, E.v., Koolman, X., Jones, A.M., 2004. Explaining income-related inequalities in doctor utilisation in Europe. Health Economics 13, 629–647.

Duggan, M., Starc, A., Vabson, B., 2016. Who benefits when the government pays more? Pass-through in the Medicare Advantage program. Journal of Public Economics 141, 50–67.

d'Uva, T.B., Jones, A.M., 2009. Health care utilisation in Europe: New evidence from the ECHP. Journal of Health Economics 28, 265–279.

Eggleston, K., 2000. Risk selection and optimal health insurance-provider payment systems. The Journal of Risk and Insurance 67, 173.

Eggleston, K., 2005. Multitasking and mixed systems for provider payment. Journal of Health Economics 24, 211–223.

Eggleston, K., Ellis, R.P., Lu, M., 2012. Risk adjustment and prevention. Canadian Journal of Economics 45, 1586–1607.

Elinder, M., Erixson, O., Öhman, M., 2023. Cognitive ability, health policy, and the dynamics of COVID-19 vaccination. Journal of Health Economics 91, 102802.

Ellegård, L.M., 2020. Effects of pay-for-performance on prescription of hypertension drugs among public and private primary care providers in Sweden. International Journal of Health Economics and Management 20, 215–218.

Ellegård, L.M., Dietrichson, J., Anell, A., 2018. Can pay-for-performance to primary care providers stimulate appropriate use of antibiotics? Health Economics 27, e39–e54.

Ellis, R.P., 1998. Creaming, skimping and dumping: Provider competition on the intensive and extensive margins. Journal of Health Economics 17, 537–555.

Ellis, R.P., Martins, B., Rose, S., 2018. Chapter 3 - Risk adjustment for health plan payment, in: McGuire, T.G., van Kleef, R.C. (Eds.), Risk Adjustment, Risk Sharing and Premium Regulation in Health Insurance Markets. Academic Press, pp. 55–104.

Galizzi, M.M., Godager, G., Li, J., Linnosmaa, I., Tammi, T., Wiesen, D., 2023. Economics of healthcare provider altruism, in: Zimmermann, K.F. (Ed.), Handbook of Labor, Human Resources and Population Economics. Springer International Publishing, Cham, pp. 1–23.

Geruso, M., Layton, T., 2020. Upcoding: Evidence from Medicare on squishy risk adjustment. Journal of Political Economy 128, 984–1026.

Geruso, M., McGuire, T.G., 2016. Tradeoffs in the design of health plan payment systems: Fit, power and balance. Journal of Health Economics 47, 1–19.

Ginja, R., Riise, J., Willage, B., Willén, A., 2022. Does your doctor matter? Doctor quality and patient outcomes. NHH Dept. of Economics Discussion Paper .

Goodman-Bacon, A., 2021. Difference-in-differences with variation in treatment timing. Journal of Econometrics 225, 254–277.

Gustafsson, P.E., Fonseca-Rodríguez, O., Castel Feced, S., San Sebastián, M., Bastos, J.L., Mosquera, P.A., 2024. A novel application of interrupted time series analysis to identify the impact of a primary health care reform on intersectional inequities in avoidable hospitalizations in the adult Swedish population. Social Science & Medicine 343, 116589.

Gutacker, N., Siciliani, L., Moscelli, G., Gravelle, H., 2016. Choice of hospital: Which type of quality matters? Journal of Health Economics 50, 230–246.

Handel, B., Hendel, I., Whinston, M.D., 2015. Equilibria in health exchanges: Adverse selection versus reclassification risk. Econometrica 83, 1261–1313.

Hennig-Schmidt, H., Selten, R., Wiesen, D., 2011. How payment systems affect physicians' provision behaviour—An experimental investigation. Journal of Health Economics 30, 637–646.

Hälso- och sjukvårdsnämnden, 2012. Protokoll hälso- och sjukvårdsnämnden 2012-05-08.

Iacus, S.M., King, G., Porro, G., 2011. Multivariate matching methods that are monotonic imbalance bounding. Journal of the American Statistical Association 106, 345–361.

Illenberger, N.A., Small, D.S., Shaw, P.A., 2020. Impact of regression to the mean on the synthetic control method: Bias and sensitivity analysis. Epidemiology 31, 815–822.

Johansson, N., Jakobsson, N., Svensson, M., 2019. Effects of primary care cost-sharing among young adults: Varying impact across income groups and gender. The European Journal of Health Economics 20, 1271–1280.

Kringos, D.S., Boerma, W.G., Hutchinson, A., Saltman, R.B., 2015. Building primary care in a changing Europe. WHO Regional Office for Europe.

Layton, T.J., 2017. Imperfect risk adjustment, risk preferences, and sorting in competitive health insurance markets. Journal of Health Economics 56, 259–280.

Lee, J.Y., Solon, G., 2011. The fragility of estimated effects of unilateral divorce laws on divorce rates. The BE Journal of Economic Analysis & Policy 11.

Lissenden, B., Balkrishnan, R., 2020. Does risk adjustment reduce vaccination in the elderly? Evidence from Medicare Advantage. Medical Care Research and Review 77, 176–186.

Lohr, S., Schochet, P.Z., Sanders, E., 2014. Partially nested randomized controlled trials in education research: A guide to design and analysis. NCER 2014-2000. National Center for Education Research .

Mackenbach, J.P., Stirbu, I., Roskam, A.J.R., Schaap, M.M., Menvielle, G., Leinsalu, M., Kunst, A.E., 2008. Socioeconomic inequalities in health in 22 European countries. New England Journal of Medicine 358, 2468–2481.

Mackenbach, J.P., Valverde, J.R., Artnik, B., Bopp, M., Brønnum-Hansen, H., Deboosere, P., Kalediene, R., Kovács, K., Leinsalu, M., Martikainen, P., et al., 2018. Trends in health inequalities in 27 European countries. Proceedings of the National Academy of Sciences 115, 6440–6445.

Malmström, M., Sundquist, J., Bajekal, M., Johansson, S.E., 1998. Indices of need and social deprivation for primary health care. Scandinavian Journal of Social Medicine 26, 124–130.

McGuire, T.G., Zink, A.L., Rose, S., 2020. Simplifying and improving the performance of risk adjustment systems. Working Paper 26736. National Bureau of Economic Research.

Mora-García, C.A., Pesec, M., Prado, A.M., 2024. The effect of primary healthcare on mortality: Evidence from Costa Rica. Journal of Health Economics 93, 102833.

Newhouse, J.P., 1996. Reimbursing health plans and health providers: Efficiency in production versus selection. Journal of Economic Literature 34, 1236–1263.

Newhouse, J.P., Price, M., McWilliams, J.M., Hsu, J., McGuire, T.G., 2015. How much favorable selection is left in Medicare Advantage? American Journal of Health Economics 1, 1–26.

Nilsson, A., Paul, A., 2018. Patient cost-sharing, socioeconomic status, and children's health care utilization. Journal of Health Economics 59, 109–124.

Nordin, M., Dackehag, M., Gerdtham, U.G., 2013. Socioeconomic inequalities in drug utilization for Sweden: Evidence from linked survey and register data. Social Science & Medicine 77, 106–117.

Oxholm, A.S., Di Guida, S., Gyrd-Hansen, D., Olsen, K.R., 2019. Taking care of high-need patients in capitation-based payment schemes–An experimental investigation into the importance of market conditions. Applied Economics 51, 5174–5184.

Paasche-Orlow, M.K., Parker, R.M., Gazmararian, J.A., Nielsen-Bohlman, L.T., Rudd, R.R., 2005. The prevalence of limited health literacy. Journal of General Internal Medicine 20, 175–184.

Riksbanken, 2021. Annual average exchange rates (aggregate). URL: https://www.riksbank.se/en-gb/statistics/search-interest-exchange-rates.

Rosano, A., Loha, C.A., Falvo, R., Van der Zee, J., Ricciardi, W., Guasticchi, G., De Belvis, A.G., 2013. The relationship between avoidable hospitalization and accessibility to primary care: A systematic review. European Journal of Public Health 23, 356–360.

Sabety, A., Gruber, J., Bae, J.Y., Sood, R., 2023. Reducing frictions in health care access: The ActionHealth-NYC experiment for undocumented immigrants. American Economic Review: Insights 5, 327–346.

San Sebastian, M., Mosquera, P.A., Ng, N., Gustafsson, P.E., 2017. Health care on equal terms? Assessing horizontal equity in health care use in northern Sweden. European Journal of Public Health 27, 637–643.

Santos, R., Gravelle, H., Propper, C., 2017. Does quality affect patient's choice of doctor? Evidence from England. Economic Journal 127, 445–494.

Schwarz, T., Schmidt, A.E., Bobek, J., Ladurner, J., 2022. Barriers to accessing health care for people with chronic conditions: A qualitative interview study. BMC Health Services Research 22, 1037.

Skovsgaard, C.V., Kristensen, T., Pulleyblank, R., Olsen, K.R., 2023. Increasing capitation in mixed remuneration schemes: Effects on service provision and process quality of care. Health Economics 32, 2477–2498.

Stabile, M., Thomson, S., 2014. The changing role of government in financing health care: An international perspective. Journal of Economic Literature 52, 480–518.

Starfield, B., Shi, L., Macinko, J., 2005. Contribution of primary care to health systems and health. Milbank Quarterly 83, 457–502.

Starfield, B., Weiner, J., Mumford, L., Steinwachs, D., 1991. Ambulatory care groups: A categorization of diagnoses for research and management. Health Services Research 26, 53.

Sundquist, K., Malmström, M., Johansson, S.E., Sundquist, J., 2003. Care Need Index, a useful tool for the distribution of primary health care resources. Journal of Epidemiology & Community Health 57, 347–352.

Sveréus, S., Kjellsson, G., Rehnberg, C., 2018. Socioeconomic distribution of GP visits following patient choice reform and differences in reimbursement models: Evidence from Sweden. Health Policy 122, 949–956.

Tange, A., Klausen, M.B., Højgaard, B., 2020. Primary health care in the Nordic countries - Comparative analysis and identification of challenges. Technical Report. VIVE-The Danish Center for Social Science Research.

Tay, A., 2003. Assessing competition in hospital care markets: The importance of accounting for quality differentiation. RAND Journal of Economics 34, 786–814.

van de Ven, W.P., Ellis, R.P., 2000. Risk adjustment in competitive health plan markets, in: Handbook of Health Economics. Elsevier. volume 1, pp. 755–845.

Vengberg, S., Fredriksson, M., Burström, B., Burström, K., Winblad, U., 2021. Money matters – Primary care providers' perceptions of payment incentives. Journal of Health Organization and Management 35, 458–474.

Weiss, M.J., Lockwood, J., McCaffrey, D.F., 2016. Estimating the standard error of the impact estimator in individually randomized trials with clustering. Journal of Research on Educational Effectiveness 9, 421–444.

Werbeck, A., Wübker, A., Ziebarth, N.R., 2021. Cream skimming by health care providers and inequality in health care access: Evidence from a randomized field experiment. Journal of Economic Behavior & Organization 188, 1325–1350.

Wolfers, J., 2006. Did unilateral divorce laws raise divorce rates? A reconciliation and new results. American Economic Review 96, 1802–1820.

Yordanov, D., Oxholm, A.S., Gyrd-Hansen, D., Pedersen, L.B., 2023. Mapping GPs' motivation — It's not all about the money: A nationwide cross-sectional survey study from Denmark. British Journal of General Practice 73, e687–e693.

Zingmark, A.M., 2013. Ersättningsmodell för vårdcentraler 2014. Beslutsunderlag Hälso- och sjukvårdsnämnden LiÖ 2013-364.

Supplementary Material to: Better Off by Risk Adjustment? Socioeconomic Disparities in Care Utilization in Sweden Following a Payment Reform

February 2, 2024

Contents

Appendix A Details on the payment scheme

This Appendix shows the capitation amounts for individuals with different characteristics in 2010-2013 and 2014-2017. All monetary values are expressed in 2017 prices.

Table A.1 displays the payment per enrolled individual for the key features of the capitation in the period 2010-2013, i.e., before the reform under study.

Table A.1: Key features of the capitation in Östergötland 2010-2013 (2017 price level)

Year	*Age-adjusted*	*Drug*	*SES*	*Geography & 75+ (11-25 km)*	*Geography & 75+ (25+ km)*
	(1)	(2)	(3)	(4)	(5)
2010	1,087	1,639	1,851	4,532	6,798
2011	1,078	1,819	1,796	4,483	6,723
2012	1,098	1,508	1,824	4,554	6,831
2013	1,142	1,483	1,874	4,679	7,018

Table A.2: Key features of the capitation in Östergötland 2014-2017 (2017 price level)

Year	*ACG*	*CNI*	*75+*	*Age-adjusted*	*Drugs*
	(1)	(2)	(3)	(4)	(5)
2014	1,163	110		792	968
2015	2,270	113		388	471
2016	3,132	237	764		
2017	3,098	239	884		

Table A.2 displays the payment per enrolled individual for the key features of the capitation in the period 2014-2017, i.e., during the phase-in period and thereafter.[1] The amounts pertain to an individual with a weight of 1.0. ACG (column 1) is a composite of three payments: payment based on diagnoses in primary care, specific prescription drugs, and general prescription drugs.

Both the ACG and the CNI payments (column 2) increased in 2016. ACG was weighted more heavily than CNI: the CNI capitation was 9% of the ACG in 2014, 5% in 2015, and 8% in 2016-2017. As a proportion of total payments, CNI was 7% in 2014-2015 and 12% in 2016-2017. In 2016, the region's calculation of CNI changed, as the parameter inhabitants 1 year or older that have recently moved to the area was removed from the index (similar to what we do in the analysis, and for the same reason).

Simultaneously with the full phase-in of ACG and CNI in 2016, Östergötland introduced a capitation for enrolled individuals 75 years or older, following a concern that the ACG did not adequately track costs for the very oldest individuals (column 3). An age-adjusted capitation (column 4) carried over from the previous period during the phase-in period, but the payment was reduced and completely phased out by 2016. Furthermore, the weighting scheme used was also changed. The age group 20-44 was still the reference group with weight 1.0 but the weights were reduced for the youngest children (0-6 years) and from 65 and up. A reduced drug capitation also carried over for two years and was fully phased out by 2016. The weighting scheme for this part of the capitation did not change much.

The area- and SES-, geography-, and age-based risk-adjustments to the capitation used during 2010-2013 disappeared with the introduction of ACG and CNI. That is, the capitation was only based on indi-

[1] The payments are for individual enrolled at a publicly owned PCCs, the payment to private PCCs was 3% higher payment for all features of the payment system to compensate for value-added tax.

vidual characteristics from 2014 and onwards, in contrast to the earlier area-based capitation, in which the payment was determined by the area the individual lived in.

Before 2014, there was a basic grant to PCCs located in towns with only one PCC. This component remained but in a changed form after 2014: the category cutoffs were reduced to below 5,000, between 5,000 and 6,500, and between 6,500 and 8,000/11,000 (depending on year). For a PCC with 5,000 enrolled patients in such a town the payment per patient increased from around 220 SEK in 2010-2013 to around 260-270 SEK in 2014-2017. For a PCC with 10,000 patients in one of these towns, the payment decreased from 80 SEK to 0 in 2014-2015, and then increased to about 110 SEK in 2016-2017 when the cutoff was increased.

Appendix B Chronic condition indicator

To classify individuals by their chronic condition status, we use the chronic condition count variable produced by the The Johns Hopkins ACG (R) System, v. 11.2.1. This variable indicates the number of diagnoses that are 'likely to last longer than twelve months and is (sic!) likely to have a negative impact on health or functional status' (Department of Health Policy and Management, 2016).

By and large, the ACG system flags diagnoses that appear on a list of chronic conditions developed by the Center for Child and Adolescent Health Policy, Mass, General Hospital for Children, in Boston, Massachusetts. According to the ACG system documentation, 'The Center for Child and Adolescent Health Policy list and the ACG System differ in definitions related to infectious diseases such as tuberculosis, peptic ulcer disease, congenital heart disease (which is generally resolved through surgical interventions at birth), gastrointestinal obstructions and perforations (likely to be acute and treatable conditions), osteomyelitis, and prematurity. These conditions are not considered chronic conditions in the ACG System chronic condition marker.' (Department of Health Policy and Management, 2016).

The technical documentation to the ACG System does not reveal the exact ICD codes of chronic conditions, but it contains a list of aggregated diagnosis categories – so-called Expanded Diagnois Clusters (EDCs) – that include the diagnoses classified as chronic conditions. Table B.1 lists these EDCs. Note that not all diagnosis codes within these EDC categories are considered chronic.

Table B.1: Chronic condition classification from the Johns Hopkins ACG (R) System

Expanded Diagnosis Cluster (EDC)	
Acute hepatitis	Hypertension, w/o major complications
Acute leukemia	Hypertension, with major complications
Acute lower respiratory tract infection	Hypothyroidism
Acute myocardial infarction	Impulse control
Acute renal failure	Inflammatory bowel disease
Acute sprains and strains	Inherited metabolic disorders
Adjustment disorder	Irritable bowel syndrome
Administrative concerns and non-specific laboratory abnormalities	Ischemic heart disease (excluding acute myocardial infarction)
Adverse events from medical/surgical procedures	Kyphoscoliosis
Age-related macular degeneration	Lactose intolerance
Anxiety, neuroses	Low back pain
Aplastic anemia	Low impact malignant neoplasms
Arthropathy	Malignant neoplasms of the skin
Asthma, w/o status asthmaticus	Malignant neoplasms, bladder
Asthma, with status asthmaticus	Malignant neoplasms, breast
Attention deficit disorder	Malignant neoplasms, cervix, uterus
Autism Spectrum Disorder	Malignant neoplasms, colorectal
Autoimmune and connective tissue diseases	Malignant neoplasms, esophagus
Benign and unspecified neoplasm	Malignant neoplasms, kidney
Bipolar disorder	Malignant neoplasms, liver and biliary tract
Blindness	Malignant neoplasms, lung
Cardiac arrhythmia	Malignant neoplasms, lymphomas
Cardiac valve disorders	Malignant neoplasms, ovary

Continued on next page

Table B.1 – continued from previous page

EDC	
Cardiomyopathy	Malignant neoplasms, pancreas
Cardiovascular disorders, other	Malignant neoplasms, prostate
Cardiovascular signs and symptoms	Malignant neoplasms, stomach
Cataract, aphakia	Migraines
Central nervous system infections	Multiple sclerosis
Cerebral palsy	Muscular dystrophy
Cerebrovascular disease	Musculoskeletal disorders, other
Chromosomal anomalies	Nephritis, nephrosis
Chronic cystic disease of the breast	Neurologic disorders, other
Chronic liver disease	Neurologic signs and symptoms
Chronic pancreatitis	Newborn Status, Complicated
Chronic renal failure	Obesity
Chronic respiratory failure	Organic brain syndrome
Chronic ulcer of the skin	Osteoporosis
Cleft lip and palate	Other endocrine disorders
Congenital anomalies of limbs, hands, and feet	Other hemolytic anemias
Congenital heart disease	Other skin disorders
Congestive heart failure	Paralytic syndromes, other
Cystic fibrosis	Parkinson's disease
Deafness, hearing loss	Peripheral neuropathy, neuritis
Deep vein thrombosis	Peripheral vascular disease
Degenerative joint disease	Personality disorders
Dementia	Prostatic hypertrophy
Delirium	Psychological disorders of childhood
Depression	Psychosexual
Developmental disorder	Psych-physiologic and somatoform disorders
Diabetic retinopathy	Pulmonary embolism
Disorders of lipid metabolism	Quadriplegia and paraplegia
Disorders of Newborn Period	Renal disorders, other
Disorders of the immune system	Respiratory disorders, other
Eating disorder	Retinal disorders (excluding diabetic retinopathy)
Emphysema, chronic bronchitis, COPD	Rheumatoid arthritis
Endometriosis	Schizophrenia and affective psychosis
ESRD	Seizure disorder
Eye, other disorders	Short stature
Failure to thrive	Sleep apnea
Fluid/electrolyte disturbances	Sickle cell disease
Gastrointestinal signs and symptoms	Spinal cord injury/disorders
Gastrointestinal/Hepatic disorders, other	Strabismus, amblyopia
Generalized atherosclerosis	Substance use
Genito-urinary disorders, other	Thrombophlebitis
Glaucoma	Tracheostomy

Continued on next page

Table B.1 – continued from previous page

EDC	
Gout	Transplant status
Hematologic disorders, other	Type 1 diabetes
Hemophilia, coagulation disorder	Type 2 diabetes
High impact malignant neoplasms	Vesicoureteral reflux
HIV, AIDS	

Appendix C SES difference in GP visits by ACG and SES

Figure C.1 shows the average number of GP visits (adjusted for age and sex) for different levels of the ACG weight, by SES. The figure shows that for most ACG weights, the SES-difference in utilization is small. At ACG levels above 1.5, which corresponds to an expected resource utilization of 1.5 times the average in the region, there is a pro-poor pattern. For lower levels, the utilization is very similar for high- and low SES individuals. Notably, due to the skewed nature of health care costs, the majority of individuals have an ACG weight below 1.5 (see Fig. 2a).

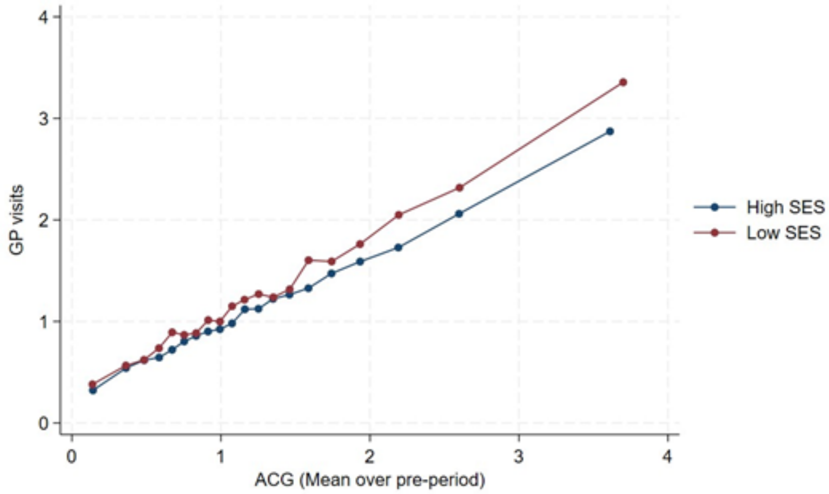


Figure C.1: GP visits by ACG level

Appendix D Synthetic difference-in-differences estimates

The DiD approach we employ relies on the untestable assumption of parallel trends. Arkhangelsky et al. (2021) suggest a synthetic difference-in-difference (SDiD) estimator that may improve the plausibility of the assumption. The SDiD estimator extends the regular synthetic control estimator, which places more weight on control group members that are similar to the treatment group, to the DiD setting. In addition to giving higher weight to control observations that are similar to the treatment group in the pre-treatment period, the SDiD estimator also assigns higher weights to pre-treatment time periods which are similar in both groups. The estimator can be implemented as a two-way-fixed effects (TWFE) model in which the squared residual is multiplied by the product of the time and unit weights.

We estimate SDiD models for the probability and number of physician and nurse visits. For computational feasibility reasons, we estimate the models on a random 30% subsample of the study population. We use the user-written Stata command *sdid* (Pailañir and Clarke, 2023). Because the command only allows us to estimate one treatment effect at a time, we estimate separate models for each of the announcement, phase-in, and post periods. I.e., we estimate three models for each of the four outcomes. Notably, within each outcome, the estimated weights do not vary across specifications, i.e., they do not depend on which post period we consider.

Table D.1: Synthetic difference-in-differences estimates

	Physician		Nurse	
	(1) Probability	(2) Number	(3) Probability	(4) Number
Announcement	-0.000382	-0.00417	0.0112***	0.0463***
	(0.00254)	(0.00434)	(0.00241)	(0.0121)
2014 reform	-0.000171	-0.00226	0.0156***	0.0886***
	(0.00273)	(0.00472)	(0.00274)	(0.0154)
2016 reform	-0.00434	-0.0158***	0.0110***	0.0807***
	(0.00281)	(0.00474)	(0.00297)	(0.0168)
N	783132	783132	783132	783132
Individuals	27969	27969	27969	27969

Note: Synthetic diff-in-diff estimates of the four outcome variables indicated by the column names. Standard errors are estimated by jackknife. * p<0.05, ** p<0.01, *** p<0.001.

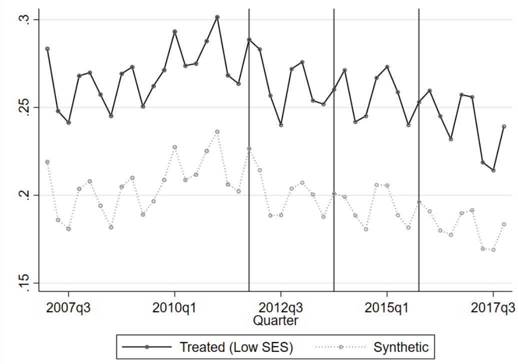
Table D.1 displays the SDiD estimates. For three of the outcomes (probability of a physician visit, number of physician visits, probability of a nurse visit), the SDiD estimates are consistent with our main results. The only difference is that the estimates of physician visits are negative and, with one exception, not significant rather than negative and significant, and the positive effect on the probability of nurse visits lasts throughout the full post period.

For the fourth outcome (number of nurse visits), the SDiD estimates are positive and statistically significant. This result stands in sharp contrast to the main specification, which yielded estimates that were negative or close to zero (Table 4).

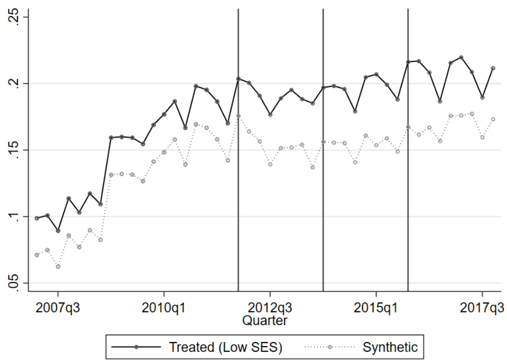
The question then is which model of nurse visits is the most credible. It is tempting to favor the SDiD, which ensures that the synthetic pre-treatment trend in the control group reproduce the pre-treatment trend of the treatment group. However, the ability to mirror pre-treatment outcomes may come at a cost. Specifically, when combining a matching method with DiD, as SDiD does, the weighting may induce a bias due to regression to the mean (RTM) in the control group. As discussed by e.g., Daw and Hatfield (2018) and Illenberger et al. (2020), RTM biases the DiD estimate when the matching method gives higher weights to control group individuals who have a true mean lower than the treatment group's but, because of random factors, have outcomes that are close to the treatment group's mean in the pre-treatment period. These extreme individuals will over time revert back to their true mean, which induces RTM bias.

The results in Daw and Hatfield (2018) and Illenberger et al. (2020) concern RTM bias in the context of matching on pre-treatment *levels* of the outcome variable. The SDiD model "reweights and matches pre-exposure trends" (Arkhangelsky et al., 2021, p. 4089), and "allows for treated and control units to be trending on entirely different levels prior to a reform of interest" (Clarke et al., 2023, p. 2). That is, the SDiD model aims to match the *trends*, and one may think that it would be immune to the type of RTM bias discussed by Daw and Hatfield (2018) and Illenberger et al. (2020). However, although the aim is to match on trends, the matching method may still give extreme control group individuals high weight in order to find well-matched trends. We show below that this seems to be the case for nurse visits but not for the other outcomes.

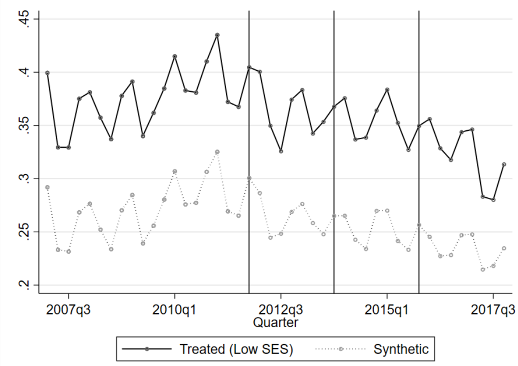
The relevance of the RTM argument for the nurse visits outcome can be grasped by considering Fig. D.1, which shows the treatment and synthetic control group outcomes for each of the four outcomes. For the outcomes in subfigures (a)-(c), the synthetic control group lies below the treatment group, at a level similar to the unweighted outcomes (see see Fig. 3 and Fig. 4 in the main text). By contrast, subfigure (d) shows that the synthetic control group has a much higher pre-treatment number of nurse visits than the treatment group. Because the treatment group had a higher average number of visits than the unweighted control group, this pattern suggests that the SDiD model gives high weights to control group individuals with extreme number of visits in the pre-treatment period. That is, exactly the problem that causes RTM bias according to Daw and Hatfield (2018) and Illenberger et al. (2020). The fact that the synthetic control group outcome starts to decline as soon as the announcement period begins further suggests that RTM may be at play – although it could just reflect a valid announcement effect.



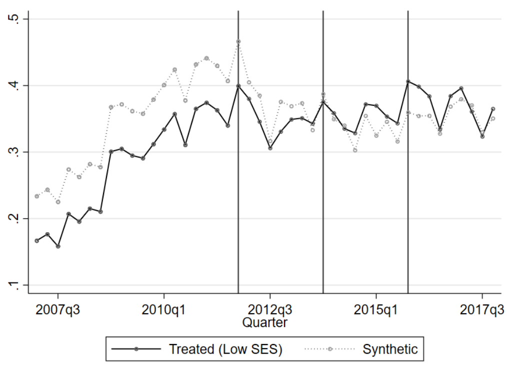
(a) Pr(physician visit)



(b) Pr(nurse visit)



(c) Number of physician visits

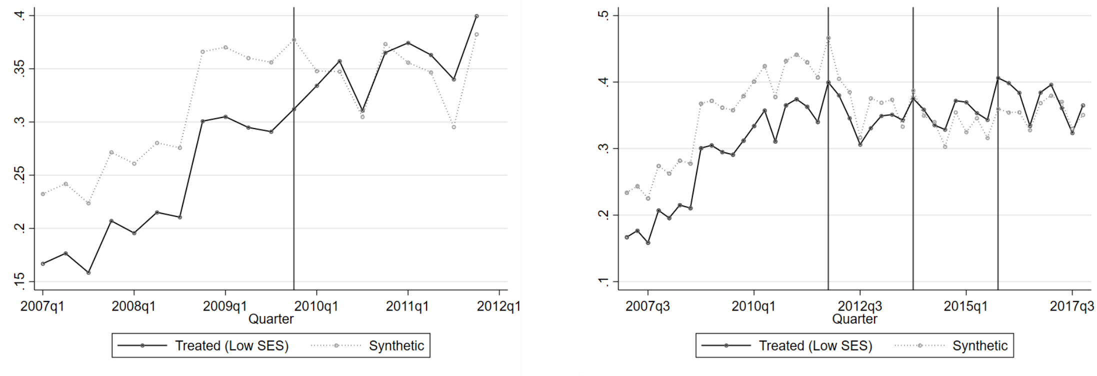


(d) Number of nurse visits

Figure D.1: Synthetic DiD results

Note: The vertical lines mark the last period of the pre-treatment period.

To further explore if RTM is the reason for the positive SDiD estimates in the model of nurse visits, we estimate another SDiD specification for the same outcome, this time using a placebo treatment period occurring earlier (2010-2011). Figure D.2(a) displays the results. Like in the previous specification (reproduced in subfigure (b) for comparison), the synthetic control group has a higher number of nurse visits than the treatment group in the (now shorter) pre-treatment period. However, as soon as the pre-treatment period is over, the synthetic outcome falls. Thus, RTM indeed appears to be a relevant concern, invalidating the SDiD model for the nurse visits outcome.



(a) Placebo treatment (2010)

(b) All post periods

Figure D.2: Synthetic DiD results for nurse visits

Note: The vertical lines mark the last period of the pre-treatment period.

For the other outcomes, the synthetic control group averages were much closer to the unweighted control group averages in the pre-treatment period. This suggests that the SDiD estimator did not have to give extreme control group individuals high weights to obtain well-matched trends, and therefore that the scope of RTM is limited. Indeed, we note that while half of the control individuals were assigned a zero weight in the nurse visits model, all control individuals were assigned positive weights, with relatively little dispersion, in the model for the other outcomes.

Appendix E Robustness of nurse visit estimates

Table E.1: Nurse visits: Difference-in-differences estimates and robustness

	Panel A: Probability of visit							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Announcement	0.00611***	0.00336	0.0118***	0.00633***	0.00528***	0.0115***	0.00611	0.00459**
	(0.00143)	(0.00207)	(0.00135)	(0.00148)	(0.00140)	(0.00142)	(0.00447)	(0.00153)
2014 reform	0.00665***	0.00232	0.0139***	0.00856***	0.00622***	0.0154***	0.00665	0.00512**
	(0.00159)	(0.00310)	(0.00150)	(0.00163)	(0.00155)	(0.00163)	(0.00353)	(0.00175)
2016 reform	-0.000807	-0.00671	0.00996***	0.00184	0.000157	0.0119***	-0.000807	-0.00233
	(0.00172)	(0.00411)	(0.00162)	(0.00177)	(0.00168)	(0.00179)	(0.00344)	(0.00189)
N	4085972	4085972	4085972	3771196	3884012	4861192	4085972	2971616
No. clusters	92863	92863	92863	85709	88273	148180	44	92863
Mean dep.	0.145	0.145	0.145	0.149	0.148	0.166	0.145	0.160
Announcement=2014 DiD	0.728	0.587	0.149	0.156	0.530	0.0100	0.748	0.728
2014 DiD = 2016 DiD	0.000	0.000	0.008	0.000	0.000	0.030	0.000	0.000
Linear trend	No	Yes	No	No	No	No	No	No
CEM	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Attritioners	Included	Included	Included	Excluded	Included	Included	Included	Included
Treatment def.	Fixed	Fixed	Fixed	Fixed	Predicted	Yearly	Fixed	Fixed
Pop. def.	Cohort	Cohort	Cohort	Cohort	Cohort	Yearly	Cohort	Cohort
Cluster	Ind	Ind	Ind	Ind	Ind	Ind	PCC	Ind
Min year	2007.000	2007.000	2007.000	2007.000	2007.000	2007.000	2007.000	2010.000
	Panel B: Number of visits							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Announcement	0.000573	-0.0159	0.0153*	-0.00247	-0.00367	0.0127	0.000573	-0.00554
	(0.00698)	(0.00880)	(0.00654)	(0.00631)	(0.00662)	(0.00736)	(0.0104)	(0.00693)
2014 reform	-0.00332	-0.0291*	0.0111	-0.000744	-0.00465	0.0145	-0.00332	-0.00944
	(0.00713)	(0.0133)	(0.00674)	(0.00725)	(0.00679)	(0.00781)	(0.00869)	(0.00756)
2016 reform	-0.0212**	-0.0564**	-0.000218	-0.0159*	-0.0209**	0.00128	-0.0212*	-0.0273**
	(0.00777)	(0.0179)	(0.00732)	(0.00797)	(0.00729)	(0.00830)	(0.00892)	(0.00836)
N	4085972	4085972	4085972	3771196	3884012	4861192	4085972	2971616
No. clusters	92863	92863	92863	85709	88273	148180	44	92863
Mean dep.	0.257	0.257	0.257	0.259	0.265	0.311	0.257	0.283
Announcement=2014 DiD	0.581	0.136	0.523	0.784	0.884	0.809	0.570	0.581
2014 DiD = 2016 DiD	0.001	0.000	0.022	0.005	0.001	0.024	0.001	0.001
Linear trend	No	Yes	No	No	No	No	No	No
CEM	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Attritioners	Included	Included	Included	Excluded	Included	Included	Included	Included
Treatment def.	Fixed	Fixed	Fixed	Fixed	Predicted	Yearly	Fixed	Fixed
Pop. def.	Cohort	Cohort	Cohort	Cohort	Cohort	Yearly	Cohort	Cohort
Cluster	Ind	Ind	Ind	Ind	Ind	Ind	PCC	Ind
Min year	2007.000	2007.000	2007.000	2007.000	2007.000	2007.000	2007.000	2010.000

Note: In each model, the dependent variable is the either the an indicator for at least one visit (Panel A) or the number of nurse visits (Panel B) at a primary care center in a quarter. The table shows three average DiDs, each computed over a number of quarterly differences-in-differences (DiD) estimates from linear regression models. The estimates contrast individuals with Care Need Index (CNI)>0 to individuals with CNI=0; Announcement shows the average of the quarterly DiD estimates in the announcement period (Q2 2012 – Q4 2013), 2014 reform shows the average of the quarterly DiD estimates after the reform (2014-2015), and 2016 reform shows the average of the quarterly DiD estimates after the price increase in 2016 (2016-2017). Announcement=2014 DiD shows p-value of test of equality of the DiDs for 2012-13 and 2014-15 periods. 2014 DiD = 2016 DiD shows p-value of test for equality of DiDs in 2014-15 and 2016-17. All models include quarter fixed effects. Individual's CNI status is measured on Dec 31 2011 unless stated otherwise (see row Treatment def). Predicted = use predicted CNI for 2011 instead of actual CNI. CEM weights balance the sample in terms of birth year and gender. Attritioners are defined as individuals moving out from the region or dying in 2013-2017. The study population is a cohort who were registered at a PCC in the region on Jan 1 2013 and lived in the region throughout 2007-2011; the exception is the column for which Pop.Def is indicated as Yearly. In that specification, the study population includes individuals registered at a PCC in the region on January 1 year t, for t=2010-2017. Standard errors are clustered by individual in all models except in column 8, in which they are clustered by the individual's PCC at baseline. * p<0.05, ** p<0.01, *** p<0.001.

Appendix F Heterogeneity by reform effect on capitation

Table F.1 shows results from DiD models of the probability and number of physician and nurse visits, estimated for two subsets of the study population.

In Panel A, the estimation sample excludes individuals who were enrolled at the two PCCs whose revenues decreased a lot due to the removal of the previous, area-based, SES adjustment – notably, the novel, CNI-based SES payment did not compensate these PCCs for the loss of area-based payment, even though they had the by far highest average CNI among PCCs in the region (.47 and .62, vs. around .15-.30 for other PCCs in 2011). In fact, from 2016 onwards, these two PCCs received a special compensation (over and above the CNI payment) that was motivated by their high SES burden.

In Panel B, the estimation sample excludes all observations with an ACG below 1. This implies that individuals only appear in the sample in years when they have above average ACG. Notably, the number of observations is approximately half of that in the main specification, which shows that below-average ACG is not a persistent phenomenon for most people.

The results from these specifications are similar to the main results (Tables 3 and 4). As these specifications in principle rule out the possibility that high-CNI (low-SES) individuals have lower capitation than they would have had under the previous payment regime, these results suggest that the main results are not driven by the fact that the payment reform reduced the capitation for some low-SES individuals.

Table F.1: Observations affected positively by the reform

Panel A: Excluding two low-SES PCCs

	(1) Pr(phys)	(2) Nr(phys)	(3) Pr(nurse)	(4) Nr(nurse)
Announcement	-0.00374*	-0.00795**	0.00168	-0.00373
	(0.00149)	(0.00255)	(0.00147)	(0.00711)
2014 reform	-0.00460**	-0.0105***	0.00395*	-0.00594
	(0.00165)	(0.00284)	(0.00167)	(0.00773)
2016 reform	-0.00995***	-0.0228***	-0.00364*	-0.0263**
	(0.00169)	(0.00285)	(0.00181)	(0.00841)
N	3842080	3842080	3842080	3842080
Individuals	87320	87320	87320	87320

"

Panel B: Excluding observations with ACG below 1

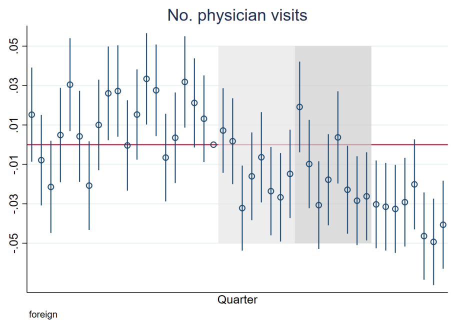
	(1) Pr(phys)	(2) Nr(phys)	(3) Pr(nurse)	(4) Nr(nurse)
Announcement	-0.00869***	-0.0188***	0.000149	-0.0134
	(0.00216)	(0.00407)	(0.00225)	(0.0123)
2014 reform	-0.0100***	-0.0227***	0.00179	-0.0211
	(0.00231)	(0.00437)	(0.00248)	(0.0121)
2016 reform	-0.0171***	-0.0408***	-0.00776**	-0.0431***
	(0.00237)	(0.00434)	(0.00261)	(0.0127)
N	1964584	1964584	1964584	1964584
Individuals	89483	89483	89483	89483

Note: The dependent variable is either an indicator for at least one visit ("Pr()") or the quarterly number of visits ("Nr") with a physician or a nurse at a primary care center. In Panel A, the sample excludes individuals who were enrolled at the two PCCs with lowest SES. In Panel B, the sample excludes observations with an ACG <1. Standard errors are clustered by individual. * p<0.05, ** p<0.01, *** p<0.001." file write myfile "
"

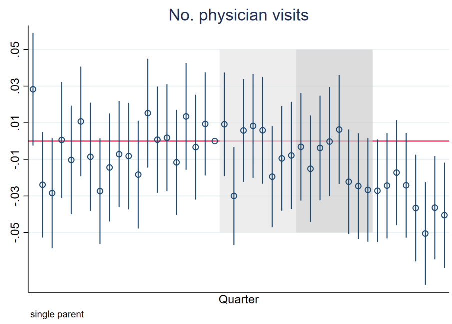
Appendix G Heterogeneity across CNI dimensions

This section shows heterogeneity estimates derived from event study models in which we contrast individuals belonging to a certain CNI category (foreign, single parent, short education, unemployed) to the whole comparison group of high-SES individuals. We estimate the effects separately for each category. Since an individual can belong to more than one CNI category (the dimensions are not mutually exclusive), it is possible to be included in more than one estimation.

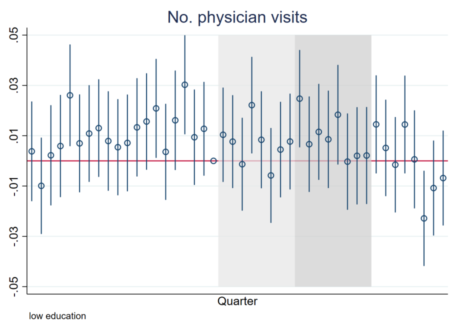
Figure G.1 indicates that there is not much heterogeneity across the CNI dimensions for physician visits. In particular, no group seem to have benefited in terms of getting more care. All groups contribute to the baseline negative estimates after 2016, and the patterns are reasonably similar before 2016.



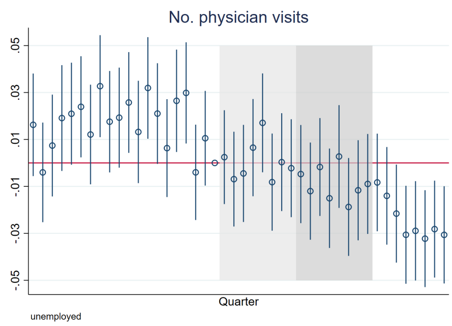
(a) Born outside EU



(b) Single parent



(c) Short education



(d) Unemployed

Figure G.1: Heterogeneity over CNI dimensions

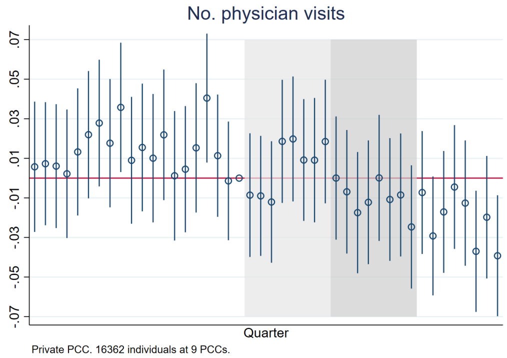
Appendix H Heterogeneity across PCC characteristics

This section examines treatment effect heterogeneity on physician visits across two types of PCC characteristics: ownership type (private/public) and the expected budget impact of the changed payment system. Figure H.1 and Figure H.2 display event-study graphs, and Table H.1 shows the estimated triple interaction terms from DiD models with standard errors clustered by the PCC.

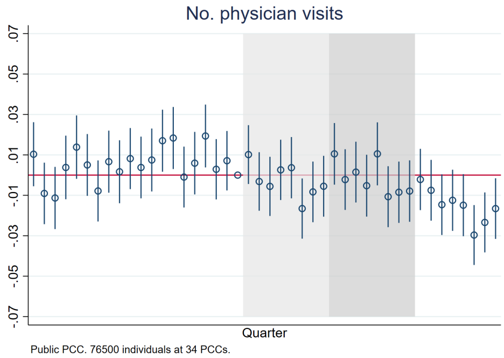
Of note, the categories of private PCCs and PCCs with expected budget increases are not capturing the same heterogeneity, two out of nine private PCCs are in the expected budget increase-category, three are in the decrease-category, and four are in the middle.

Figure H.1 shows that the negative treatment effects in 2016-2017 are visible for both private and public PCCs. The estimates are larger and the trend starts somewhat earlier for private PCCs, albeit the estimates are noisy for this type. The DiD estimate in Table H.1 indicate that the difference between private and public PCCs is significant in both post-periods.

The less positive estimates for treated individuals registered at private PCCs may indicate that private units are less prone to spend the additional money received for high CNI individuals to provide additional care for that group. However, it is important to acknowledge that the results do not by themselves show that the money was directly channeled as profits; for instance, the private PCCs might have been spent the funds on other groups, or on longer consultations for treated individuals.



(a) Private PCCs



(b) Public PCCs

Figure H.1: Heterogeneity over ownership type

Note: The dependent variable is the quarterly number of physician visits at a PCC. Separate models for individuals registered at private (upper panel) and public (lower panel) PCCs. The shaded areas represent the announcement period and the first two years of the post-reform period.

In Figure H.2, we group PCCs according to whether they were projected to experience large increases or declines in their budget in 2014, i.e., the first post-reform year. To produce the projections in these budgets, the health care administrators applied 2014 payment rules to the characteristics of the listed population in August 2013. Thus, the information represents the best guess at the end of 2013 of what the financial situation would look like in 2014.[2]

We divide PCCs into three groups defined by how large a budget shock they experienced in 2014: a more than 2% decrease ("Decreasers" in the table; 25 PCCs), a more than 2% increase ("Increasers"; 6

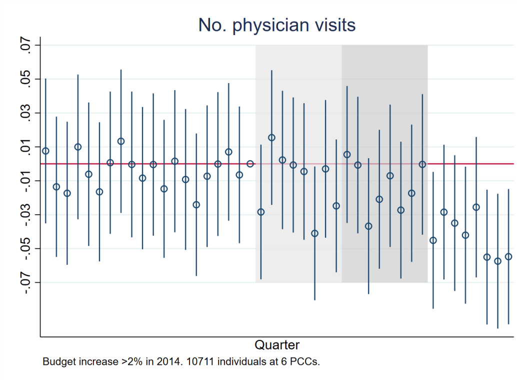
[2] The listed population in August 2013 could to some extent be affected by announcement effects of the reform, but the announcement is unlikely to be a major influence. We also have access to budgets projected using data from April 2013; the deviation from the August projection is 4% or below for all but three PCCs.

PCCs), or smaller budget changes than that (*"Stable"*; 12 PCCs). The reason why as many as 25 PCCs are classified as expecting a large budget decrease relates directly to the transfer of home care to the municipalities, which implied an overall budget cut of around 5%.[3]

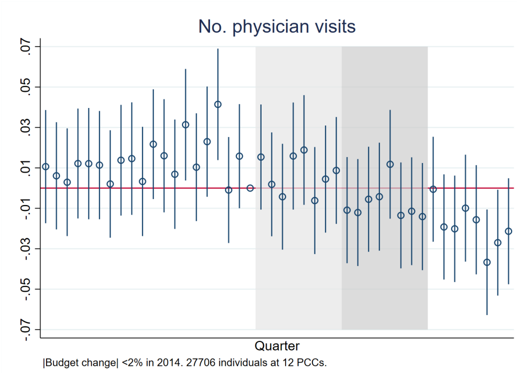
Figure H.2 indicates that the overall pattern of negative estimates, especially in the 2016-2017, is present across the three categories but perhaps clearest among Winner PCCs. The DiD estimate in Table H.1 indicate that the difference between PCCs with expected budget decreases and PCCs with stable budgets is significant in both post-periods, and that the difference between PCCs with expected budget decreases and increases is significant in 2016-2017.

The pattern of heterogeneity across PCCs with expected budget increases and decreases provides evidence against the contemporaneous nurse reform influencing our estimates. The nurse reform was an important determinant of the expected budget change and there is a small over-representation of treated individuals enrolled at PCCs with expected budget decreases: The proportions of treated individuals are 21%, 20%, and 25% at PCCS with expected increases, stable budgets, and decreases, respectively. The corresponding proportions for all low-SES individuals are 25%, 25%, and 32%. If the nurse reform had a differential impact on our treatment and comparison groups, we would expect a stronger negative association with visits for PCCs with decreasing budgets because they were more affected by the nurse reform and have a larger proportion of treated individuals enrolled. As mentioned above, this pattern is not what we observe.

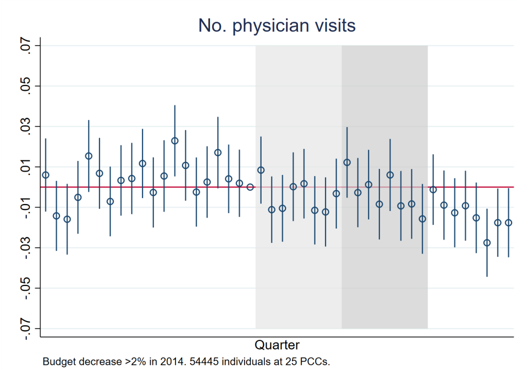
[3] As noted previously, the budget cut was accompanied with a transfer of district nurse capacity to the municipalities, reflecting the narrower scope for the PCC services.



(a) Increasers



(b) Stable



(c) Decreasers

Figure H.2: Heterogeneity over expected budget change in 2014

Note: The dependent variable is the quarterly number of physician visits at a PCC. Separate models for individuals registered at a PCC with a more than 2% budget increase (upper panel), more than 2% budget decrease (lower panel), or a change within that range (middle panel) in 2014 compared to 2013. The shaded areas represent the announcement period and the first two years of the post-reform period.

Table H.1: Provider-level heterogeneity in DiD estimate on physician visits

	(1)	(2)
Announcement × Middle	0.00213	
	(0.00641)	
Announcement × Winner	0.00167	
	(0.00790)	
2014 × Middle	-0.0139*	
	(0.00680)	
2014 × Winner	-0.00252	
	(0.00945)	
2016 × Middle	-0.0146*	
	(0.00699)	
2016 × Winner	-0.0217*	
	(0.00906)	
Announcement × private		0.000351
		(0.00731)
2014 × private		-0.0167*
		(0.00689)
2016 × private		-0.0138*
		(0.00607)
N	4085972	4085928
Clusters	44	43

Note: Estimates of heterogeneity in DiD estimates over provider groups defined by (1) projected budget change in 2014 (2) private/public ownership. In all models, the dependent variables is the quarterly number of physician visits. The estimates shown in the table are triple interaction terms of the form *treatment group* × *period* × *provider group*; i.e., *treatment group* indicates individuals with high CNI (low SES), *period* is either the announcement period, the first part of the post reform period (2014-2015), or the period after the price increase (2016-2017), and *provider group* indicates either PCCs with projected budget change in the mid-range (Middle = at most a 2% change) or a large increase (Winner = more than 2%) (columns 1-2), or private PCC (columns 3-4). All models include individual and year fixed effects, and use CEM weights to balance the sample in terms of birth year and gender. Standard errors are clustered by PCC. * p<0.05, ** p<0.01, *** p<0.001.

Appendix I Additional results for the time-varying population

Figure I.1 shows event-study estimates for the time-varying population discussed in section 6.1. The DiD estimates for this population, reported in column 7 of Table 3, were positive. However, the event-study estimates in Figure I.1 indicate that there is an upward-sloping trend in the pre-treatment period.[4] When we include a linear trend estimated on the pre-treatment quarters in our DiD specification (i.e., according to Eq. 2), the positive estimates disappear: the estimate for the announcement period is then -0.0026 ($p = 0.478$), the estimate for the phase-in period is -0.0036 ($p = 0.534$), and the estimate for the final period is -0.016 ($p = 0.035$). Thus, there is no strong evidence of increases in physician visits for this study population either.

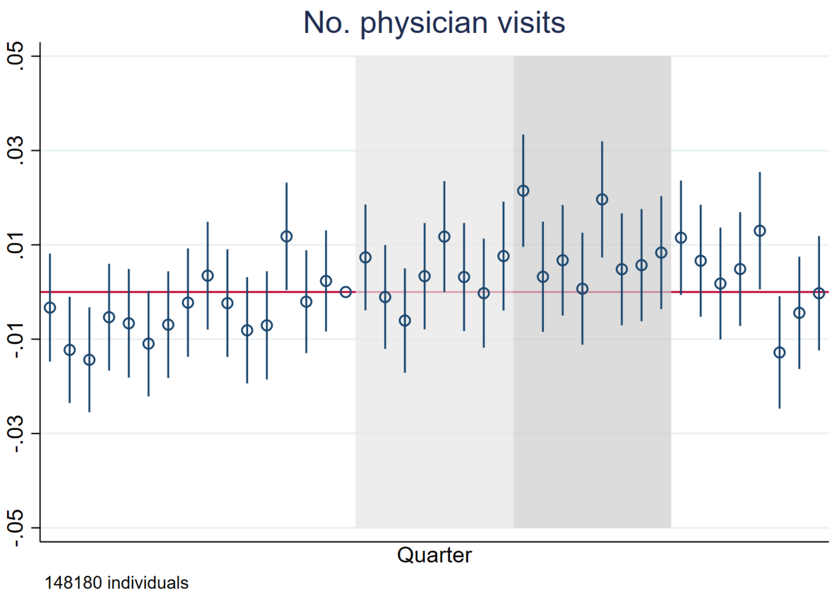


Figure I.1: Event study for time-varying population (2008-2017)

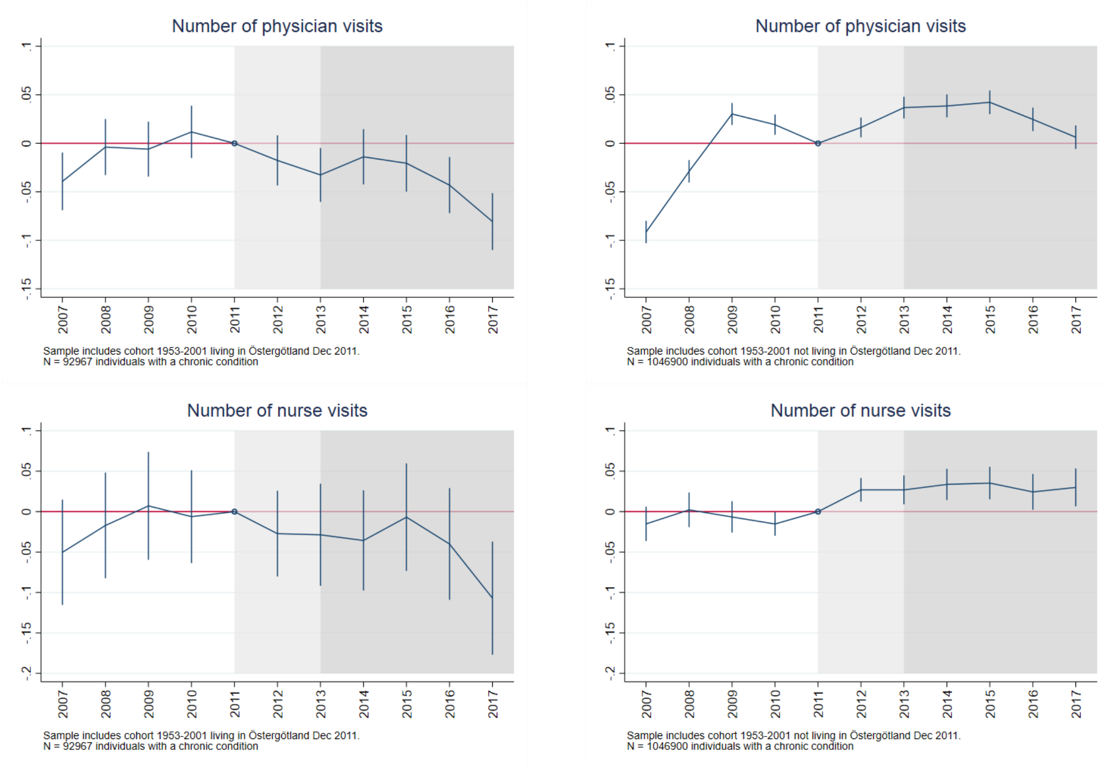
[4]A potential reason for the upward-sloping trend is that the composition of this sample may have changed due to the relatively large and increasing immigration of refugees from the Middle East and Africa during this period (see e.g., migrationsverket.se/Om-Migrationsverket/Statistik/Beviljade-uppehallstillstand-oversikter.html), who would be included in the treatment group and may have needed more care. As our primary sample is not time-varying, it avoids this issue.

Appendix J Comparison with other regions

In this section, we estimate event-study specifications for three other regions from which we have access to data on physician and nurse visits, ACG weights, and secondary care outcomes: Stockholm, Västra Götaland, and Skåne (the three largest Swedish regions). Note that Stockholm introduced major changes to its payment system in 2016, including new risk-adjustment models similar to the ones used in Östergötland and already in place in Västra Götaland and Skåne from 2009.

We report the results in three figures. The left panels in each figure show estimates for a sample mimicking the main study population in Östergötland and the right panels shows estimates for a combined sample of the other three regions. The high-CNI definition used in these estimations is almost as in the main estimations; the exception is that individuals who moved within a region in 2010-2011 are included in the high-CNI group. We cannot separate out these individuals from the high-CNI group in the other three regions as we lack data on who moved within these region. (As mentioned, including this group leaves our main estimates in Östergötland virtually unchanged.) Furthermore, we use yearly averages for all variables, for computational reasons.

Figure J.1 shows estimates for the number of physician (top panel) and nurse visits (bottom panel). For both outcomes, the difference between the high and low-CNI groups is trending slightly downward in Östergötland, and the difference is significant in 2016-2017 for physician visits, and in 2017 for nurse visits. In the other regions, both outcomes show increases in the post-period. For physician visits, there is however an increasing but strongly non-linear trend in the pre-period, which is likely driven by the large changes to the market structure in these regions around the introduction of their patient choice reforms (2007 in Stockholm, and 2009 in Skåne and Västra Götaland). Östergötland was comparatively less affected by for example entry of private providers after its patient choice reform (also in 2009). The trend in the pre-period is flatter for nurse visits while there is a slight uptick in 2012, which lasts throughout the post-period. Importantly, none of the outcomes display a continuously growing difference in visits, which we would expect if the trends were driven by diverging health trends.

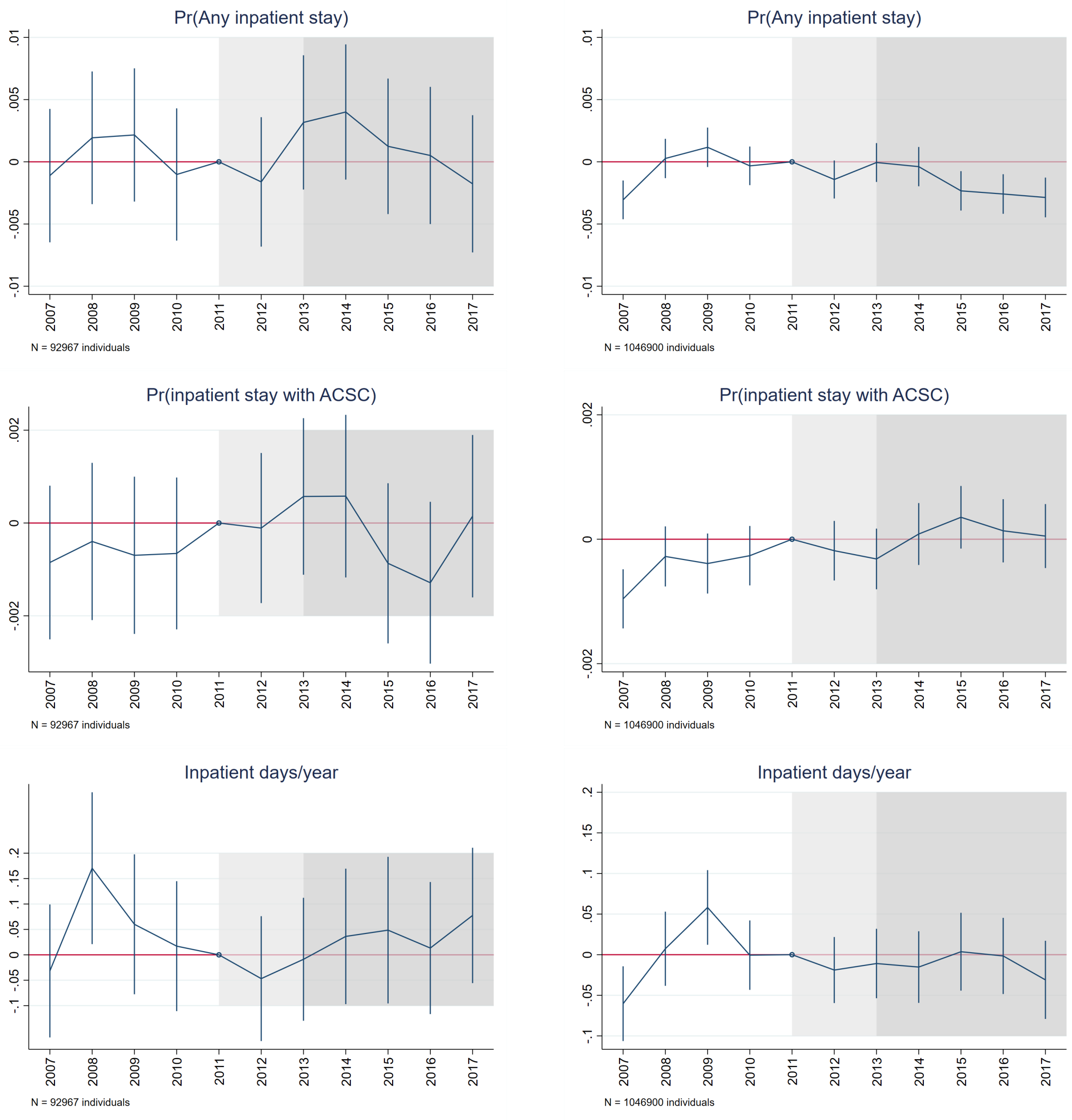


(a) Östergötland (b) Other regions

Figure J.1: Physician and nurse visits, comparison with Stockholm, Västra Götaland and Skåne regions; annual data 2007-2017

Note: The estimates are for individuals who lived in Östergötland (a) or any of the other three regions (b) at the end of 2006-2012. The outcomes are gathered from the primary care registers from the four regions. The treatment definition is almost as in the main estimations; the only exception is that individuals who moved within a region in 2010-2011 are included in the high-CNI group. The estimates are weighted to make the low and high-CNI groups balanced in terms of birth year and gender.

Figure J.2 shows estimates for the probability of an inpatient stay, of an ACSC hospitalization, and the number of inpatient days per year. The figure shows that the high CNI/low SES group in Östergötland had a roughly similar development on these variables as the other three regions, except that the probability of an inpatient stay appears to have increased slightly more in Östergötland in 2013-2014. The lack of positive "post-reform" trends in the other regions indicates that over study periods of this length, diverging health trends between low and high SES individuals need not be an issue.



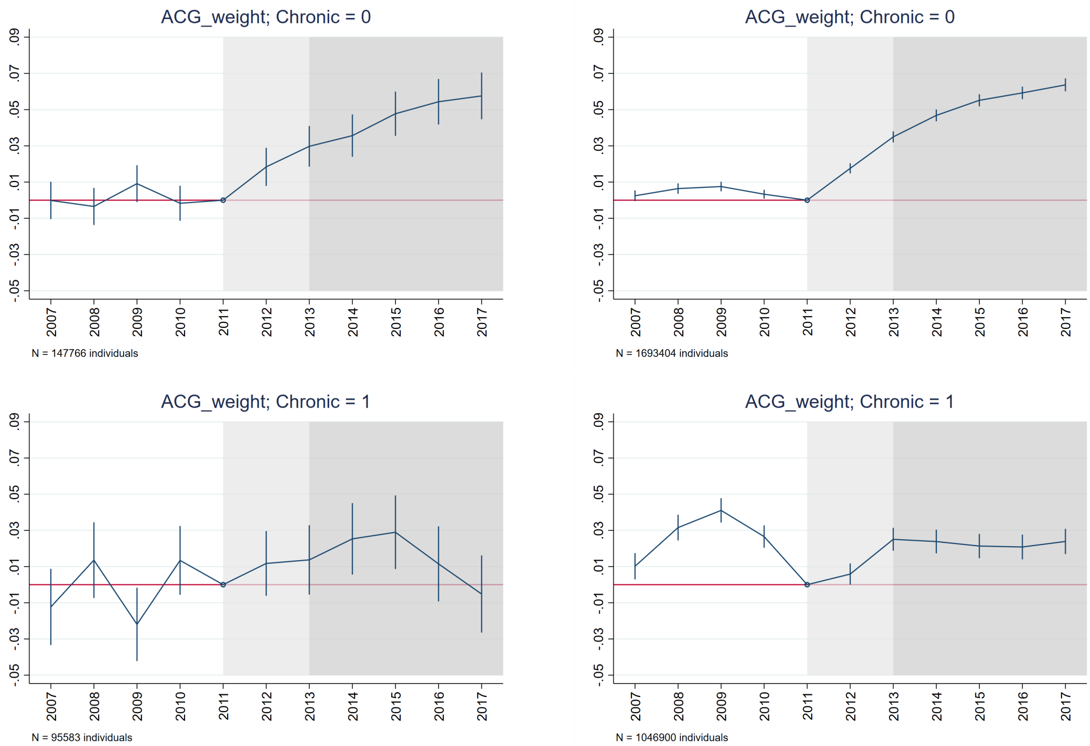
(a) Östergötland (b) Other regions

Figure J.2: Hospital outcomes comparison with Stockholm, Västra Götaland and Skåne regions; annual data 2007-2017

Note: The estimates are for individuals with a pre-period chronic condition who lived in Östergötland (a) or in any of the other regions (b) at the end of 2006-2012. The outcomes are from the national inpatient register. The treatment definition is almost as in the main estimations; the only exception is that individuals who moved within a region in 2010-2011 are included in the high CNI group. The estimates are weighted to match the high CNI group in terms of birthyear and gender.

Figure J.3 shows event study graphs for the ACG weight in Östergötland and the three other regions. Here, we show results both for individual *without* a pre-existing chronic condition (upper figures) and for the group with such a condition (lower figures). The figures for individuals without a chronic condition indicate a lack of pre-trends and thereafter a sharp increase in all regions. The reason is likely that the definition of the group without chronic conditions forces the trends to be similar in the pre-period, where after the latent higher probability of becoming chronically ill in the low SES group is allowed to manifest itself. This implies that it is not credible to attribute the increase in Östergötland to the payment reform. Therefore, we do not study other outcomes for this group.

For the group with chronic conditions, the pre-trends in the other three regions are not parallel, most likely due to the patient choice reforms implemented in those regions in 2008-2009 (including the introduction of ACG compensation in Västra Götaland and Skåne). However, there is little evidence of a steady divergence between the low- and high-SES groups when it comes to the ACG weight, neither across the whole period, nor in the first and second post-period in these regions. As was the case for the secondary care outcomes, these estimates do not provide strong evidence of diverging health trends between low and high SES individuals with chronic conditions.



(a) Östergötland

(b) Other regions

Figure J.3: ACG, comparison with Stockholm, Västra Götaland and Skåne regions; annual data 2007-2017

Note: The estimates are for individuals who lived in Östergötland (a) or any of the other three regions (b) at the end of 2006-2012. The upper (lower) figures are for individuals without (with) a pre-period chronic condition. The treatment definition is almost as in the main estimations; the only exception is that individuals who moved within a region in 2010-2011 are included in the high-CNI group. The estimates are weighted to make the low and high-CNI groups balanced in terms of birth year and gender.

References

Arkhangelsky, D., Athey, S., Hirshberg, D.A., Imbens, G.W., Wager, S., 2021. Synthetic Difference-in-Differences. American Economic Review 111, 4088–4118. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20190159, doi:10.1257/aer.20190159.

Clarke, D., Pailañir, D., Athey, S., Imbens, G., 2023. Synthetic difference in differences estimation. arXiv preprint arXiv:2301.11859 .

Daw, J.R., Hatfield, L.A., 2018. Matching and regression to the mean in difference-in-differences analysis. Health Services Research 53, 4138–4156.

Department of Health Policy and Management, 2016. The Johns Hopkins ACG System Version 11.1 Technical Reference Guide , Bloomberg School of Public Health, Johns Hopkins University.

Illenberger, N.A., Small, D.S., Shaw, P.A., 2020. Impact of Regression to the Mean on the Synthetic Control Method: Bias and Sensitivity Analysis. Epidemiology 31, 815–822. URL: https://journals.lww.com/10.1097/EDE.0000000000001252, doi:10.1097/EDE.0000000000001252.

Pailañir, D., Clarke, D., 2023. SDID: Stata module to perform synthetic difference-in-differences estimation, inference, and visualization. Statistical Software Components URL: https://ideas.repec.org//c/boc/bocode/s459058.html. publisher: Boston College Department of Economics.