

The implications of reducing activity-based payments for emergency care*

This briefing paper summarises the findings of the ESHCRU2 project 3.1 Analysis of purchaser-provider contracts: modelling risk sharing and incentive implications. The project has focused on the implications of payment reform of what is termed blended payment for emergency care.

**Martin Chalkley, Hugh Gravelle, Nikita Jacob, Rita Santos,
and Luigi Siciliani
September 2023**

*This project is funded by the National Institute for Health and Care Research (NIHR) Policy Research Programme (reference PR-PRU-1217-20301). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.



Summary

- Reducing activity based payments is one key component of the move to blended payment for emergency care and an underlying principle for Integrated Care Systems in England. It results in a weaker link between activity and payment and therefore, according to the usual concept, weakens the incentive to engage in that activity.
- Whilst reducing the inappropriate use of hospitals for emergency admissions is widely held to be necessary and desirable, the interplay of incentives in relation to emergency care suggests that there might be unintended consequences from this kind of policy. Furthermore, the impact of a change in payment will vary from one local health system to another. There is thus a need to understand both the extent and the drivers of differences across England.
- Our research has focused on understanding these issues. We have constructed a conceptual framework that allows for the co-determination of hospital admissions and the attendances at Accident and Emergency Departments (AEDs) that give rise to these. Our approach is founded upon the idea that both providers (hospitals) who receive payment and purchasers (CCGs) who make those payments make decisions that depend on one another. In this setting changing a payment system changes incentives on both sides and the interaction between the resulting choices determines the outcome in terms of the emergency care system.
- One key implication is that the performance of an emergency care system – the combination of attendances and admissions – needs to be considered as a whole and whilst payment system reform typically targets admissions, there are likely to be consequences for attendances. We show that this will generally be in the direction of a reduction in admissions being accompanied by an increase in attendances. Hence, the movement away from activity-based payment encounters a potential trade-off between different elements of hospital-based emergency services.
- Whilst theory provides a useful organising framework, understanding the implications of payment reform in practice requires empirical evidence. Therefore, our research has focused on the variation between hospitals and CCGs in terms of their (respectively) admission and attendance rates holding as many 'other things equal' as possible.
- We find there is considerable variation across both groups.

- High admitting hospitals have around three times the admission rates of low admitting hospitals – after accounting for as many other factors as we can find relevant data on. These variations are statistically significant and are distributed unevenly across the country.
- The different admission rates are associated with different expected responsiveness of admissions to payment reform. Hence knowledge of these differences would seem essential in planning and advising on the implementation of payment reform.
- The responsiveness of admission rates to varying activity prices is likely to be small – our calibration of the empirical results to the theoretical model suggests that a 10% reduction in price will result in at most a 2 percentage point reduction in admissions, but in most circumstances the reduction will be much less.
- In terms of CCGs there is a similar ‘other things equal’ variation in attendance rates at AEDs by their constituent populations. High attendance CCGs exhibit approximately three times more attendances than low attendance CCGs.
- Once again, this variation is not evenly distributed across the country and suggests that policy implementation may need to differ across locations.
- We do not have sufficient evidence to determine the potential responsiveness of attendance rates to changing activity prices.
- Putting the two components together there are local systems that exhibit both high attendance rates and admissions. These ‘hot’ systems might be particularly relevant for further study because our theoretical framework indicates that payment reform cannot improve performance on both measures simultaneously.

Introduction

Until recently, under Payment by Results (PbR) and the National Tariff Payment System (NTPS), hospitals in England were paid according to the number and type of patients they treat. In this system payment reflects activity since the greater the throughput of patients the more money a hospital will receive. Under an approach that has been called blended payment which applied from 2019 for emergency admissions, and which is intended to apply to an increasing range of hospital services, the role of these activity-based payments is being reduced. This trend towards reducing activity-based payment is an underpinning principle of the newly emerging Integrated Care Systems (ICS).

In this paper, we interpret and summarise the findings of our research into the implications of adopting blended payment – and thereby reducing activity-based payments - for emergency admissions.

We have produced three research papers focusing on both the conceptual (theoretical) and empirical evidence for this change. Whilst we are reporting here on the specifics of emergency care some of our findings translate readily for other services for which there is a move away from activity-based payment.

In the next section we set out the key elements of our approach and the relevant features of emergency care that informed our research. The following three sections then set out in more detail the basis for our findings - each section corresponds to one of the academic research papers we have produced. The subsequent section brings the various elements together to show the relevance of our findings for understanding how different areas of England can be expected to have substantially different implications of a move towards reducing activity payments.



Key elements of our approach

Blended payment through the lens of economics

The focus of economics is on the role of mechanisms that influence decisions and thereby guide the use of resources. From this perspective blended payment can be seen as a movement away from the use of prices for emergency care. An essential element of reducing activity based payment is to reduce the price a hospital will receive for each additional patient that is treated.

There is much in the policy documents describing blended payment that is concerned with the specifics of how, and under what circumstances lower prices will apply, but from an economics perspective these are largely distracting details: blended payment corresponds to lower prices and an offsetting payment of a fixed sum.

Whether in the context of consumers choosing which goods to buy or firms deciding where to invest in the future prices play a crucial role. Economists view prices as important determinants of behaviour. In healthcare settings this role is often described as one of providing incentives.

Hence, our first approach as economists is to view blended payment as reducing the incentive to engage in activity. This

accords well with the rationale for introducing blended payment in the context of emergency admissions to hospitals, where it has been suggested there might otherwise be a tendency for hospitals to over-admit patients.

Our engagement with policy colleagues however suggested another, and much less commented-on role of prices within the NHS. Just as a lower price for the recipient represents a lower incentive to 'earn' that payment, for the payer it lowers the incentive to avoid the payment. This observation turns out to be important in conceptualising the impact of blended payment for emergency care and we have made it central to our research.

Whereas much of economics in health care focuses on incentives for providers, we have balanced our approach by examining the implications for purchasers. The results of considering the bilateral incentives of both purchasers and providers are so fundamental that they constitute a large part of what we feel we have contributed to knowledge through our research.

To understand this requires the consideration of the particular context that has been the focus of our research.

The emergency care setting and its key features for our analysis

Much of the focus on emergency care has been on admissions to hospital. This is understandable given the high costs of inpatient care. However this focus risks distracting from the sources and drivers of attendances at Accident and Emergency Departments (AEDs) from which admissions arise, and from the point of view of considering system costs that distraction is problematic if measures to reduce admissions have the effect of increasing AED attendances disproportionately more.

The existing literature on attendances is relatively scant and tends to focus on the characteristics of individuals that give rise to an increased propensity to attend an AED. However, from our discussions with stakeholders it is clear that there are many aspects of the service provision that can be expected to influence attendances – and these may be subject to influence by various agencies within the NHS, specifically purchasers. At the time of our

study, these were Clinical Commissioning Groups (CCGs).

This raises an important observation. A CCG is the purchaser of hospital services so it has an inherent incentive to avoid admissions under a pricing system in which it is responsible for paying hospitals. But a CCG also makes decisions that impact on attendances at AED and the lower the price they pay for subsequent emergency admissions, the less they may be inclined to avoid those attendances that give rise to admissions.

This insight – that there is a potential linkage between payment for admissions and the effort that goes into avoiding attendances at AEDs – is noted by the agencies involved (hospitals and CCGs) but has not previously featured in consideration of payment incentives. It forms a key feature of the emergency care process and its dependency on payment models that we took as our point of departure.

The three outputs and their interrelationships

Resolving the interdependence between admissions and attendances and the interplay between two decision making agencies (hospitals and CCGs) requires original and innovative modelling. We developed that model in our first output (Chalkley et al., 2022a).

The recognition that the motivation and priorities of both hospitals and CCGs affects the impact of payment reform and

guided our empirical strategy for outputs 2 and 3 (Chalkley et al., 2022b, 2023).

Output 2 examines variations in hospital behaviour in respect of admissions and Output 3 is focused on variation across CCGs in respect of the attendance of their populations at AEDs.

Whereas it is commonplace to refer to a theoretical model when conducting empirical studies, it is comparatively rare

to closely link the two. Often this is because the goals of theory and empirical investigations are quite disparate.

For our project however we have aligned the goals of all of our research to focus on the impact of adopting blended payment

for emergency care. Thus our theoretical framework is bespoke (although it is capable of being generalised) and it was used as the foundation for empirical study. This is an example of constructing an approach to data that is directed by theory rather than simply 'guided' by it.

Theoretical framework

The purpose of the theoretical framework is to provide a means for thinking about the determinants and interdependencies of decisions surrounding emergency care, and in particular to understand how different payment mechanisms will impact on outcomes. We broadly follow an approach that is commonplace in economic analysis of incentives – principal-agent theory. In traditional principal-agent theory the nature of payment between a purchaser and provider of services is usually viewed as being designed by the purchaser to ensure that the provider has an incentive to act appropriately. For the reasons set out above we depart from that approach by assuming that both purchaser and provider – henceforth CCG and hospital – are motivated by the payment system to make decisions that (a) impact on each other and (b) respond to the payment system that is in place between them.

This suggests a framework in which the payment system is taken as given so that we examine the choices of a hospital and CCG in response to it. We model both the

hospital's choices and actions given the payment system and the actions of the CCG and the CCG's choices given the payment system. Each party's decisions are the best response to the decisions of the other, given the payment regime. This is commonly described as a Nash equilibrium and enables us to examine the effect of changing the payment regime on the jointly dependent choices made by the hospital and CCG.

We give specific context to our model by assuming that hospitals choose what proportion of patients to admit, whilst CCGs take decisions which affect the number of attendances at emergency departments (AEDs) by their populations. The model shows how admissions and attendances depend on contextual factors (the motivation and costs of hospital and CCGs, the payment system, the extent and variability of underlying demand for emergency care) and resolving the interdependence of admissions and attendances. Our model can be understood by reference to Figure 1.

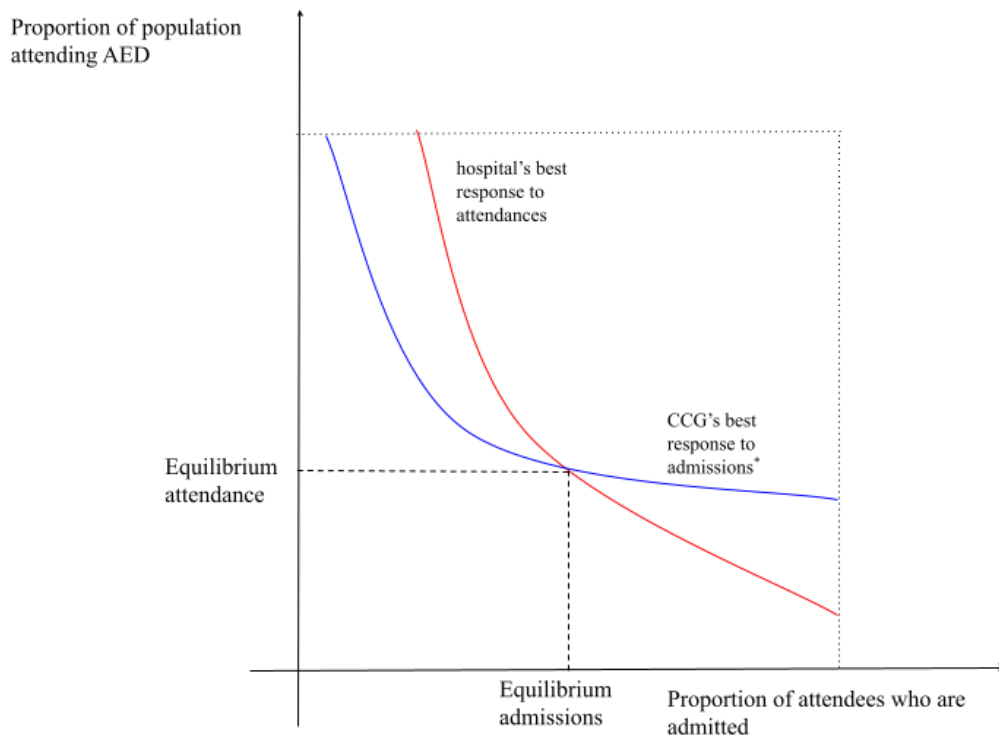


Figure 1: Purchaser and provider best responses

The outcome in terms of admissions (the proportion of patients who attend an AED who will be admitted) and attendances (the proportion of a CCG’s population that will attend an AED) lies at the intersection of the two curves in the diagram. The location and shape of each curve is influenced by the specific circumstances of the hospital (for the red curve) or the CCG (the blue curve). Both curves will also be influenced by the payment system and that is the focus of our analysis.

We can show that as the price of admissions reduces (the movement to

blended payment) the red curve will shift leftwards, whilst the blue curve will shift rightwards. This implies that whilst admissions (for a given level of attendances) will reduce, attendances will rise.

This model also provides a means for interpreting different outcomes across different purchaser and provider pairs. As shown in Figure 2, the equilibrium attendance and admissions for each pair can be located in one of the four quadrants.

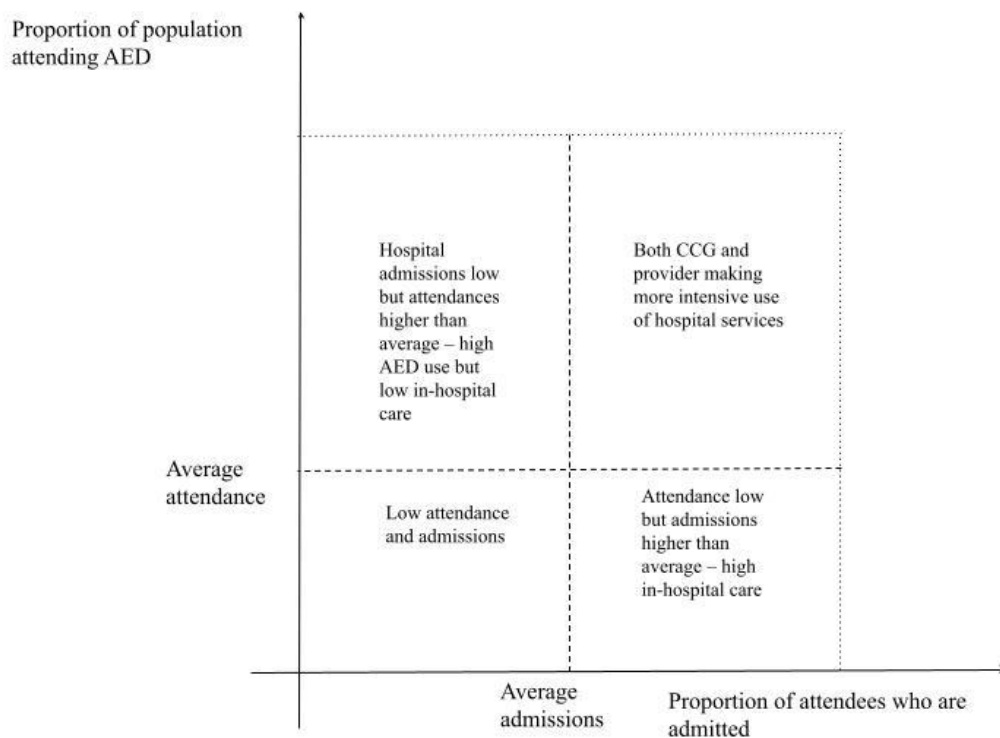


Figure 2: Interpretation of purchasers and providers pairs equilibrium attendances and admissions

In assessing whether there are risks associated with reducing activity-based payments it is important to know in which quadrant a particular CCG-hospital pair is

operating. In particular, given the risk of increasing attendances, the upper two quadrants in the figure might be viewed as being problematic.

Policy implications

- The experience of any one local health economy (the mix of attendances and admissions) will vary according to the specific circumstances of the hospital and CCG.
- The possible configuration of admissions and attendances that can be achieved through varying prices of admissions and attendances will therefore similarly vary across local health economies.
- Hence the impact of blended payment needs to be thought of in terms of the circumstances and behaviour of hospitals and CCGs.
- We established that a desired configuration (in the sense of maximising welfare) may not be achievable from varying prices alone, thus suggesting that policies in addition to blended payment may be required.

Variation in admissions across hospital providers

The theoretical framework suggests that differences between hospitals, or between CCGs, will give rise to different outcomes in terms of admissions and attendances – and will also influence how effective (or not) changes in payment will be in influencing those outcomes. We focused first on variation in admission rates across hospitals. There is rich data concerning which patients who attend AED are subsequently admitted to inpatient care, so the first task was to match data across different sources in order to facilitate analysis.

Considering each individual AED attendance, we then sought a model that establishes the specific influence of a hospital's decisions to admit that patient. We use a linear regression analysis and incorporate hospital fixed effects to capture the potential influence of hospital decisions. We relate this empirical model

closely to the underlying theoretical framework set out in Output 1, such that we can interpret the hospital fixed effects in terms of its underlying decision process and the influences on that. We then combine theory and empirical evidence by calibrating our theoretical model with our empirical findings and conduct some assessment of the likely sensitivity of admissions to variation in payment through a blended payment system.

Our key findings concern the variation in admissions across hospitals having accounted for as many differences in the patient and the environment from which they come as possible. We are able to present graphical representation of this variation in admission rates 'other things equal'. The results can usefully be summarised in two figures.



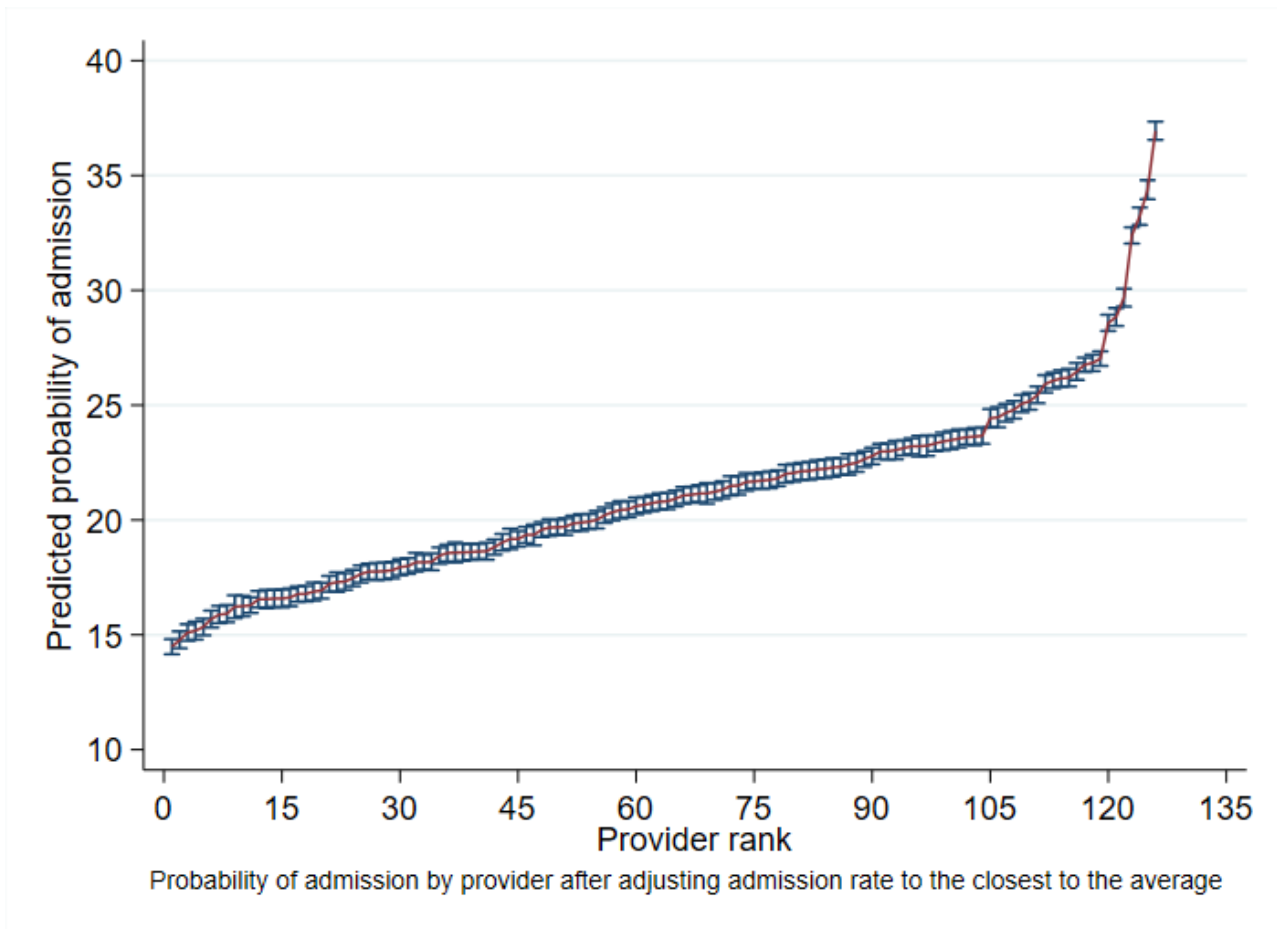


Figure 3 - Providers adjusted emergency admission rate

In Figure 3 above, hospitals are ranked according to their propensity to admit – other things equal i.e. after correcting for case-mix, patient characteristics, mode and time of arrival etc.

As can be seen these admission probabilities range from around 2% for low admitting hospitals to around 34% for high admitting hospitals. This illustrates considerable variation.

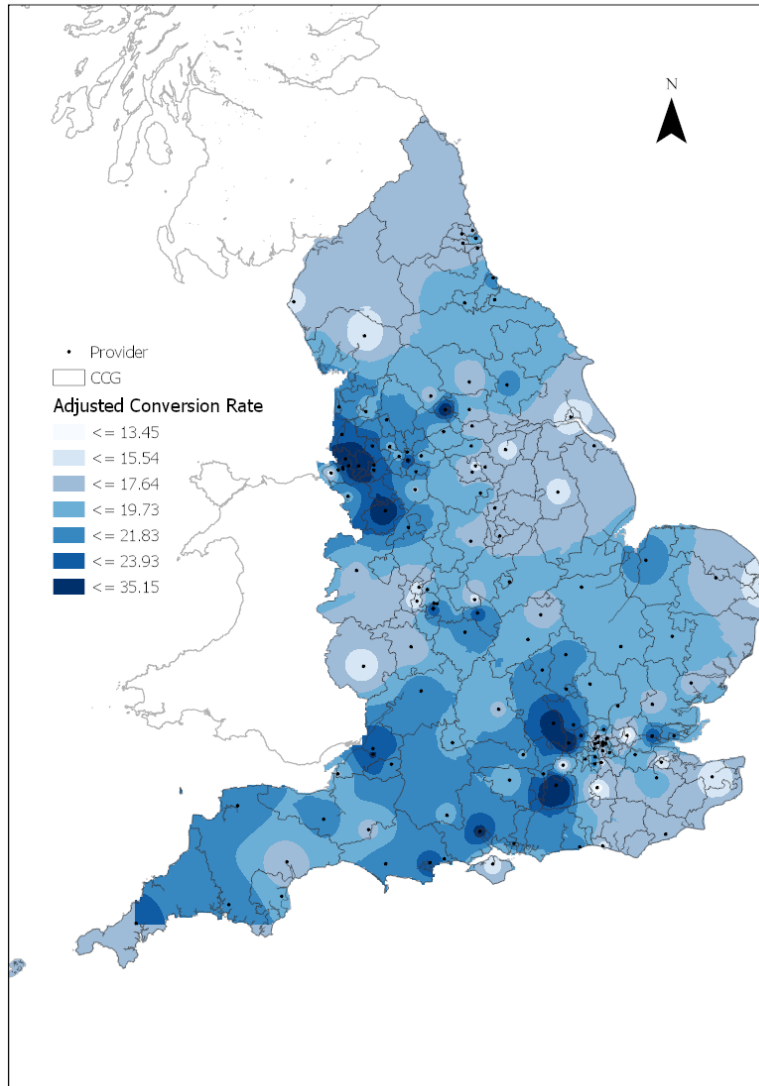


Figure 4 - Heat map of providers adjusted emergency admission rate

Figure 4 shows how this variation does not occur randomly over locations. There are local health economies that exhibit higher than average admissions (after correcting for patient characteristics) in the South West, North West and western central regions.

Relating this variation back to our theoretical model we establish that the likely responsiveness of admissions to prices is small - very large reductions in price are consistent with only very small reductions in admissions.

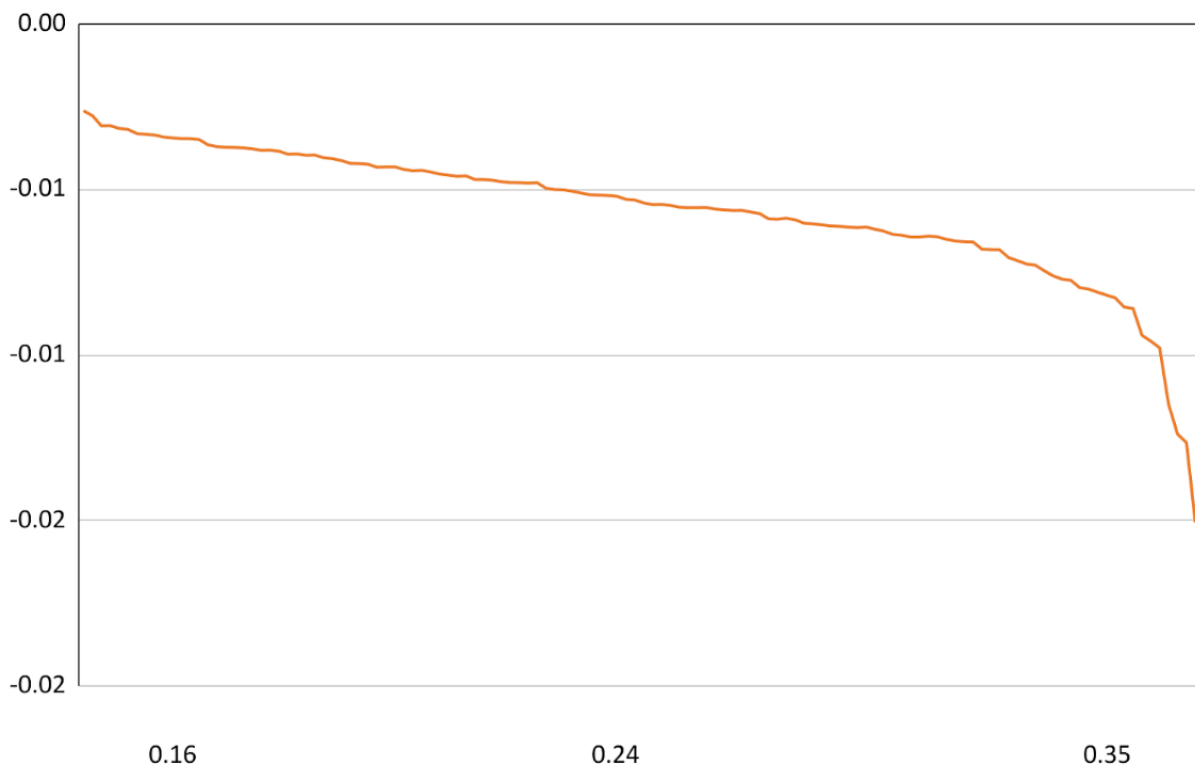


Figure 5 - Impact of 10% decrease on emergency admission price on adjusted conversion rate (vertical axis) ranked by their initial conversion rate (horizontal axis)

Figure 5 shows the responsiveness of hospitals, with different activity volumes, to a 10% price decrease, “other things equal”.

However, this likely responsiveness does vary substantially across hospitals – by a factor of 5:1.

Policy implications

- These results indicate that payment reform has different implications and a different imperative across different locations.
- In some instances, there are already very low admissions, and hence there would appear to be little requirement for payment reform to further influence this. Furthermore, in these same instances the impact of payment reform is likely to be very small.
- There is a need to understand better how very large differences between hospital admissions occur even after patient characteristics and other factors have been accounted for. This would seem to be an important general requirement for policy intervention in this domain, as the evidence strongly indicates that one size will not fit all.

Variation in attendances across CCGs

As in the consideration of differences between providers, our framework indicates that variation across CCGs will give rise to different outcomes in terms of both attendances and admissions. Relative to admissions, attendances have been little studied and our approach is novel in that we seek to examine variation in population admission rates across CCGs in the English NHS.

In contrast to our study of admissions, where we utilised data at the level of individual patients, in this study we model the admission rate of the population of a GP practice. To estimate the contribution of a CCG we include in this model a CCG fixed effect. As with Output 2, the fixed effects can be interpreted as the

idiosyncratic contribution of a CCG after external factors have been accounted for.

We again used linear regression and included control variables reflecting both the characteristics of the underlying population, the GP practice's characteristics and accessibility to minor emergency services. The findings can be related to the theoretical framework set out in Output 1 and we go on to consider the combination of (hospital specific) admissions and (CCG specific) attendances.

As with admissions our key findings are in relation to the magnitude and variation of CCGs' influences on attendances and can usefully be summarised in Figures 6 and 7.



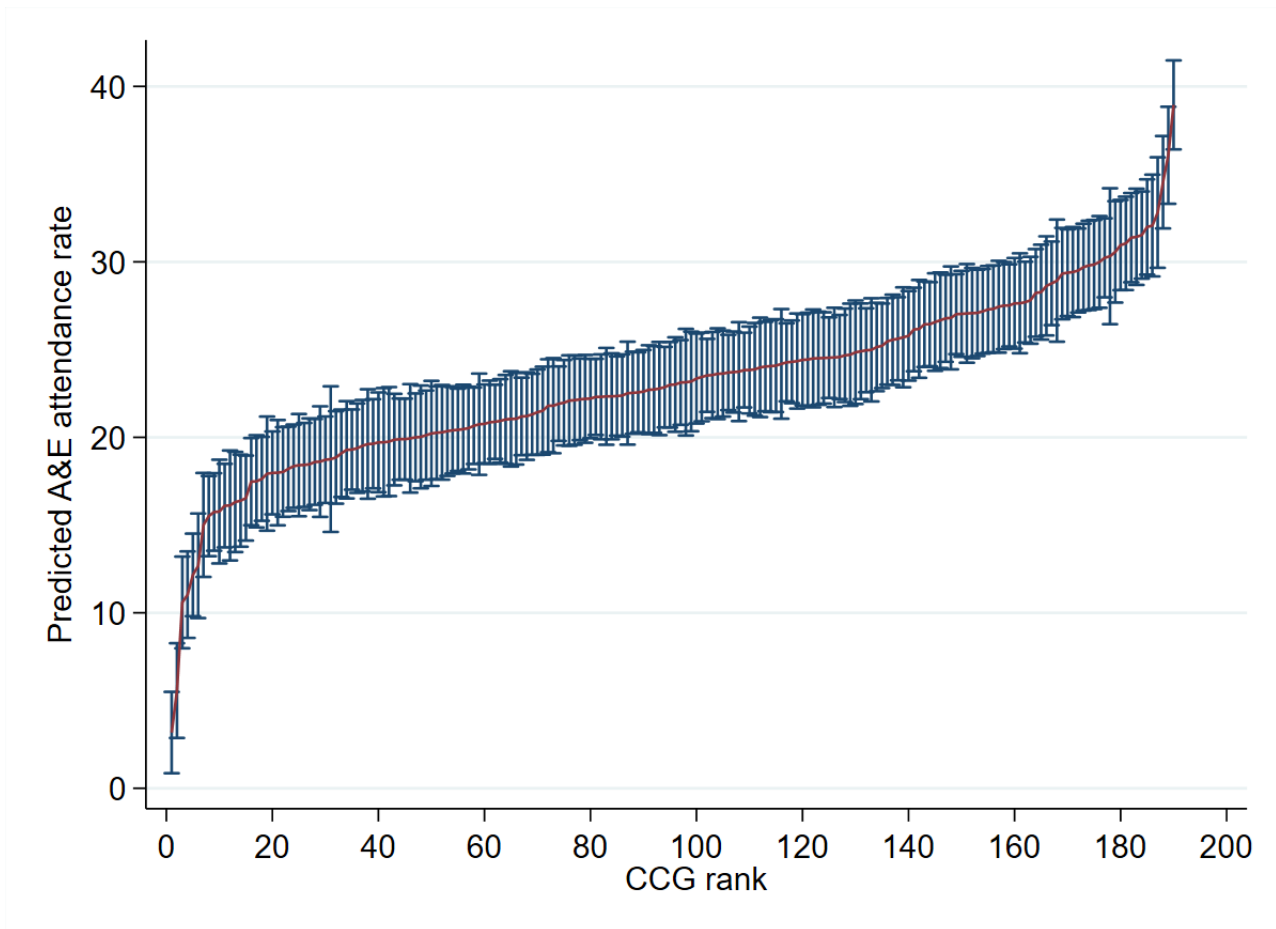


Figure 6 - Adjusted A&E attendance rate across CCGs

In this figure CCGs are ranked along the horizontal axis according to their attendance rates having accounted for differences in their populations and the nature of GP provision in their location. The relatively larger confidence intervals (denoted by the extent of the 'whiskers' in the plot) follows from the coarser information that we have available. Whereas the admissions data was at the level of individuals here we aggregate up from the level of GP practices. Nevertheless, the differences between CCGs are substantial and statistically

significant. Low attendance CCGs have around 10% of their population attending an AED whilst for high attendance CCGs the figure is more than 30%.

Once again, this variation across CCGs is not evenly distributed across England (Figure 7).

CCGs with higher than average admissions after correcting for patient and GP characteristics (the darker areas in the figure) are concentrated in the North West, North East and South East regions.

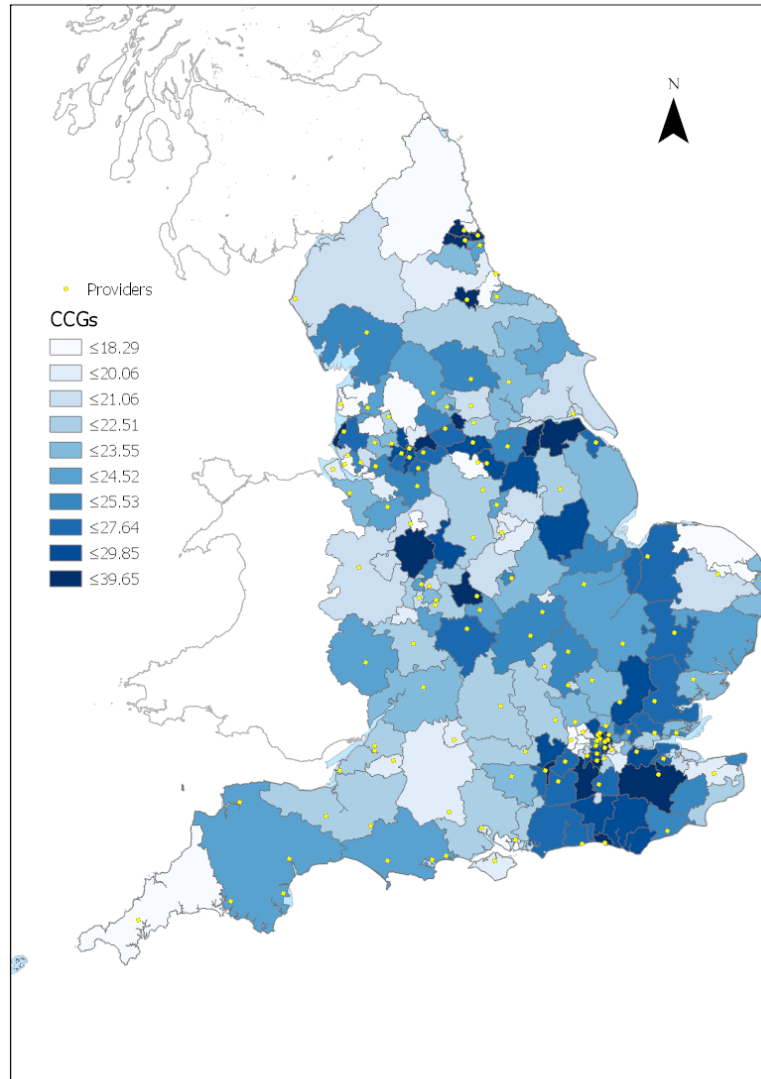


Figure 7 - Spatial distribution of CCGs adjusted A&E attendance rate

Policy implications

- Variation in attendance at AEDs that can be attributed to specific CCGs has not previously been examined. Our results suggest that this variation is substantial and that it is important to understand it more, especially in the context of implementing payment reform.
- There are instances where high attendance rates might be made higher still by the implementation of blended payment which suggests that caution should be exercised in rolling out a uniform policy.

Emergency care as a system

Drawing together the previous two sections, we can consider the emergency care system outcome as the consequence of an interaction between a hospital and a CCG. In many cases of course a single CCG will contract with multiple hospitals and a single hospital may serve many CCGs. However, if we restrict attention to those circumstances where there is a dominant hospital for a CCG – which we interpret as a single trust being the destination for 85% or more of a CCGs attendance at AED and where those

attendances account for 50% or more of that trusts AED activity – then we can combine results from our empirical provider and CCG models.

This serves to illustrate how much combined variation there is in both attendances and admissions, again after accounting for the different nature of populations and the differences in areas regarding GP provision. The results are shown in the following Figures 8 and 9.



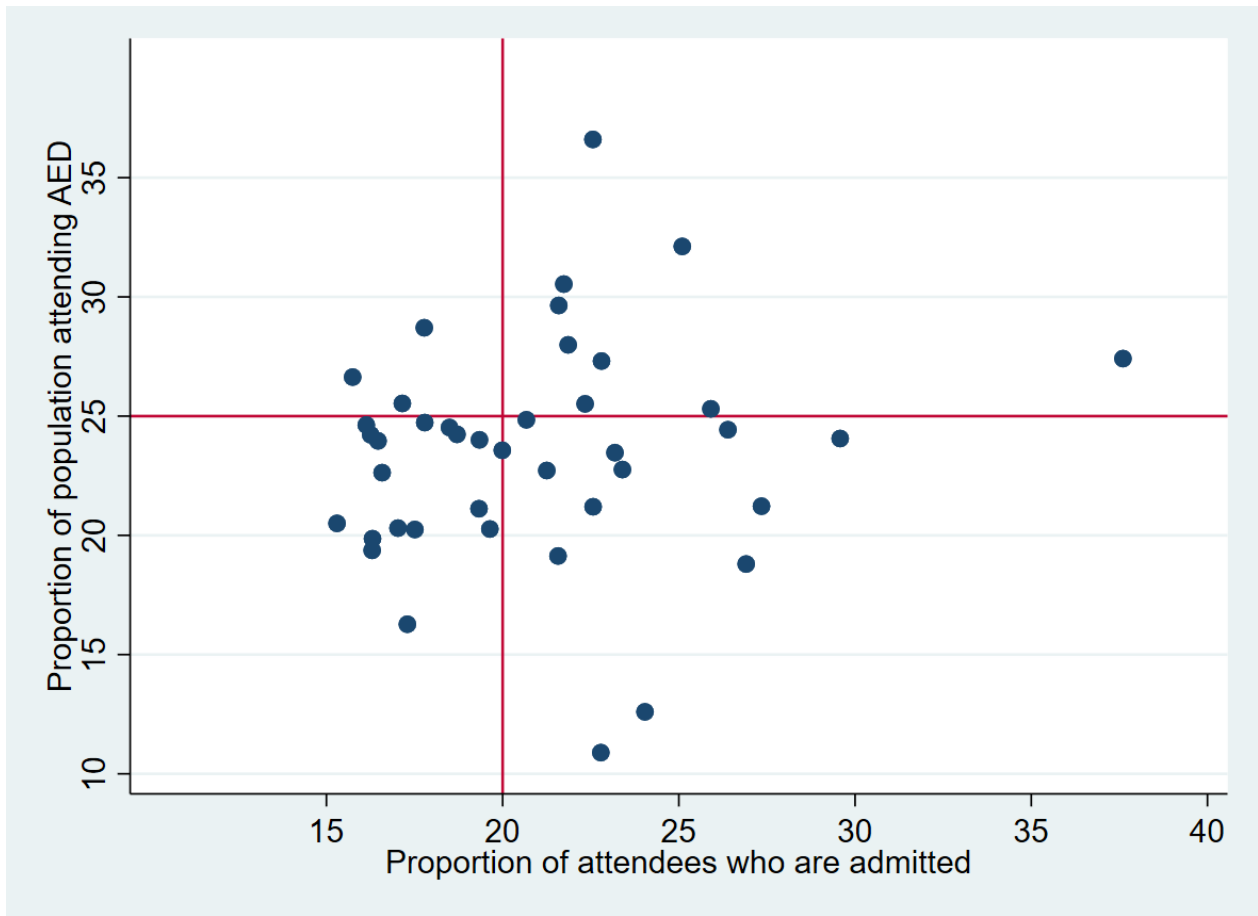


Figure 8 - Scatter plot of hospital-CCG pairs of A&E attendance and admission rates

Figure 8 is the empirical analog of Figure 2 in Section 3 and shows the dispersion of local health systems in terms of their admissions and attendances. For example, points in the top right quadrant represent health systems that make above average use of hospital services in terms of both admissions and attendances. In colloquial terms these are systems that *run hot*. Our theoretical analysis suggests that whilst implementing payment reform to moderate the price of admissions will reduce admissions it may increase

attendances, so that these local systems will likely remain hot. Furthermore, depending on the relative costs and capacity pressures in those areas the movement towards fewer admissions at the cost of higher attendances may be detrimental.

As with the component parts (of admissions and attendances) the geographical dispersion of hot systems is not random.

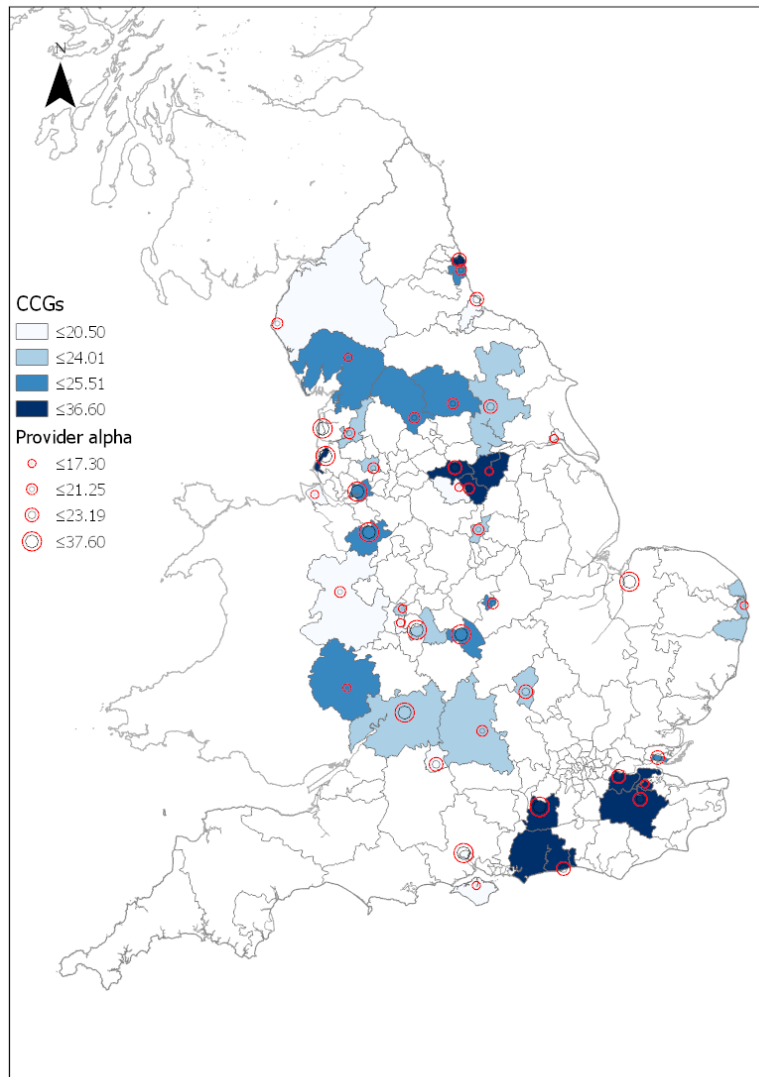


Figure 9 - Spatial distribution of hospital-CCG pairs of A&E attendances and admission rates.

In Figure 9, dark areas represent high attendance CCGs whilst large red circles indicate high admission hospitals.

The hottest spots are in the South East and northern central areas. More general hotspots exist in the North and Midlands.

Policy implications

- The results indicate that the payment reform will have different implications across the local health economies (CCGs and hospital pairs).
- There is a need to understand better why some local health economies are 'running hot' with high levels of attendances and admissions.
- Impact of blended payment needs to be implemented with additional policies to maximise patients' welfare, especially on 'run hot' local health economies.

References

Payment reform, purchaser and provider decisions and the performance of emergency healthcare systems: The case of blended payment in the English NHS • Chalkley M, Gravelle H, Jacob N, Santos R, Siciliani L • 2022a • CHE Research Paper 187 • [Download Publication](#) • [Download Lay Summary](#)

The Potential for Payment Reform to Influence Emergency Admissions: The Case of Blended Payment in the English NHS • Chalkley M, Gravelle H, Jacob N, Santos R, Siciliani L • 2022b • CHE Research Paper 188 • [Download Publication](#) • [Download Lay Summary](#)

The role of payment reform in influencing Accident and Emergency Department attendances: variation across Clinical Commissioning Groups • Chalkley M, Gravelle H, Jacob N, Santos R, Siciliani L • 2023 • CHE Research Paper 191 • [Download Publication](#) • [Download Lay Summary](#)

