Dutch shared savings program targeted at primary care: reduced expenditures in its first year

Hayen, A.; Berg, M.J. van den; Struijs, J.N.; Westert, G.P.

Health Policy 125 (2021) 489–494

Contents lists available at ScienceDirect

Health Policy

journal homepage: www.elsevier.com/locate/healthpol

Dutch shared savings program targeted at primary care: Reduced expenditures in its first year

Arthur Hayen[a,*], Michael Jack van den Berg[b], Jeroen Nathan Struijs[b], Gerard Pieter Westert (Gert)[c]

[a] Tilburg University, PO Box 90153, 5000 LE, Tilburg, the Netherlands
[b] National Institute for Public Health and the Environment, PO Box 1, 3720 BA Bilthoven, the Netherlands
[c] Radboud University (Radboud University Medical Center), PO Box 9101, huispost 114, 6500 HB Nijmegen, the Netherlands

ARTICLE INFO

Article history:
Received 7 October 2019
Received in revised form 16 October 2020
Accepted 26 January 2021

Keywords:
Payment reform
Shared savings
Primary care

ABSTRACT

In countries where GPs fulfill a central role in the health care system, like in the Netherlands, the lack of value-based incentives in GP payment systems may have negative consequences for value delivered in other parts of the health care spectrum. We evaluate an experiment in which GPs were allowed to share in savings in total health care expenditures, conditionally on achieving quality targets. At least in theory, these so-called 'shared savings contracts' incentivize GPs to become critical gatekeepers, coordinate the provision of care and substitute for specialist services when appropriate. This study evaluates a Dutch shared savings program targeted at GPs. This study employs a difference-in-differences design using a regional control group of non-participating GPs. We find that program participation led to savings in health care expenditures (-2%), while patient satisfaction was unaffected and while the results for other quality indicators were ambiguous. Additional analyses show that savings have been predominantly realized by lowering the volume of specialist care, and that almost every participating GP displayed cost-saving behavior. This finding suggests that shared savings contracts, even when added as a mere complemented to existing volume-based payment models, already elicit substantive effort to increase the value of health care provided.

1. Introduction

Traditional payment models and their financial incentives, like fee-for-service, are typically ill-aligned with the overall aim of increasing patient value [1,2]. In several countries, GPs fulfill a central role in the health care system [3], implying that when value-based incentives are lacking at this level, this has important consequences for value delivered in other parts of the health care spectrum.

In the Dutch health care system, GPs are gatekeepers to specialist services, coordinate the provision of chronic care services, and may substitute for particular hospital services. A truly value-based incentive is one that incentivizes GPs to internalize the value implications of their referral policy and decisions on substitution of services – which would otherwise be fully carried by the health insurer. This incentive is currently lacking (see Appendix A, containing a brief description of the GP payment model).

In order to add this financial incentives to the current payment model, Menzis, a Dutch health insurer, started experimenting with a 'shared savings contract' [4–8]. Under a shared savings contract, GPs are held accountable for the medical spending and quality of services their patients receive across the full continuum of care (including services by out-of-network providers) and share in expenditure savings when they outperform an expenditure benchmark. The extent to which they share in savings depends on their score on a number of quality metrics. As they are now being held financially accountable for costs and quality across the entire care spectrum, GPs would be incentivized to internalize the value implications of their decisions.

Menzis first piloted the shared savings contract with 'Arts en Zorg' (AEZ), a national chain of primary care centers (henceforth referred to as 'GPs') [9]. The program was implemented in July 2014, and is still operative. In this study, we present the first-year results. Our research question is two-fold: (1) did the shared sav-

* Corresponding author. Present address: Leiden University Medical Center, PO Box 9500, 2300 RA Leiden, the Netherlands.
E-mail addresses: a.p.hayen@lumc.nl (A. Hayen), michael.vandenberg@oecd.org (M.J. van den Berg), jeroen.struijs@rivm.nl (J.N. Struijs), gert.westert@radboudumc.nl (G.P. Westert (Gert)).

https://doi.org/10.1016/j.healthpol.2021.01.013

ings arrangement affect medical spending and quality, and (2) have these changes been instrumental in achieving value?

2. Background

2.1. The 'shared savings' pilot

Appendix A contains a general description of GP care in the Netherlands, and a more detailed description of the shared savings contract

At the time of study, the GP payment model consisted of quarterly capitation fees for each person enrolled with the practice, and small fee-for-service payments for patient visits and specific medical procedures. GPs also received bundled payments for providing a predefined bundle of chronic care services. In the pilot, a shared savings contract complemented the traditional GP payment system. Under the contract, participating GPs agreed to take accountability for medical spending and quality of care within their population of listed patients, and were held accountable for the full set of services covered by Menzis's insurance policies, also when these services had been delivered by other providers.

In the contract, accountability was promoted through a one-sided shared-savings model, in which GPs could earn part of the savings they generated in the entire spectrum of health care. Expenditure savings were evaluated against an expenditure benchmark, consisting of a three-year weighted average of GPs' own historical expenditures, trended forward by the growth rate in health care expenditures of patients from non-participating GPs in the same city. Both parts of the benchmark were corrected for case mix differences, such that all underlying patient populations reflected the GP's current population. As health care expenditures typically portray high variability across insured, a savings threshold was added, using the basic principles of statistical testing. In case a GP passed this savings threshold, its entire result qualified as a saving [10]. The resulting payout to GPs was a function of both absolute performance and improvement on performance, on a wide range of quality indicators (with better performance corresponding to higher pay-outs) (see Appendix B), the sharing rate and a payout maximum. For reasons of confidentiality no details can be given on the exact size of the payout maximum, but it is to be understood as a small, incremental, payment, at most comprising a small percentage of GP revenue (similar to those reported in other programs [10]). The contract also included a maximum amount of investment costs GPs could book to the program. A detailed description of the program, including the pros and cons of alternative design choices, can be found elsewhere [9].

The shared savings contract poses several new financial risks to both the health insurer and the GP. Even though the shared savings risk is one-sided, meaning that a GP only shares in savings and not in losses, a GP might be unable to recoup its investment costs in case of losses. As for the health insurer, even though the difference between GP and benchmark spending needs to be statistically significantly lower than zero, it could still be the case that a similar result would have been achieved in the absence of the program. Also, because the savings result is calculated on 'gross' expenditures (the sum of actual claims), a health insurer is 'paying twice' in case, for example, a participating GP lowers hospital spending in a hospital that was promised a budget. After all, these savings are shared with the GPs but the hospital still receives the full budget – as negotiated. For these hospitals, health insurers have to lower budgets in the next year, but ultimately, whether this number can be recouped depends on the insurer's bargaining power. The health insurer still decided to use 'gross expenditures', as these better reflect actual care delivery. In case the insurer used 'net expenditures', the GPs would have also been at risk of the health insurer's purchasing strategy, for which they cannot be held accountable (and this would reduce the incentive's strength [11]).

In the Netherlands, GPs routinely receive data on (benchmarked) performance, and are also able to ask the health insurer for additional data. This did not change during the pilot. However, in response to the pilot, the GPs did make use of this possibility, and asked for aggregated data on hospital use, lab tests and treatment patterns associated with high-cost patients. Due to their central role in the Dutch health care system (Appendix A), GPs have a large number of tools to lower total health care spending. Interventions observed during the pilot included, among others, the use of electronic nudges to prescribe drugs efficiently, interprofessional consultations in case of doubts as to whether a patient should be referred to hospitals and substitution of small, surgical, procedures.

3. Materials and methods

3.1. Study population

We obtained Menzis claims and enrollment data from 2011 to mid 2015, for patients continuously enrolled with Menzis, and registered with either a GP from the intervention group or with a GP from a control group (total sample: 22 GPs, 25,560 patients; intervention group: 8 GPs, 9974 patients). As a control group, we used a random sample of insured living in the same city of Enschede, who were registered with a non-participating GP. A local comparison group ensures that the market conditions are similar for both groups of GPs (e.g. hospital prices), which are, in part, influenced by the insurer's regional health care purchasing teams. GPs from either group that were not continuously operative during the period of study were excluded from the analysis (1 GP, 73 patients). We also excluded people who switched GPs during our period of study (427 patients) [19]. Our final sample consisted of 25,060 patients and 21 GPs (intervention group: 7 GPs, 9690 patients).

3.2. Study variables

We analyzed individual quarterly spending and enrollment. For each insured person in each quarter, we summed all insurance claims that were covered by the contract. In order to limit the effect of outliers, we capped quarterly spending at the 95th percentile. We created separate spending categories based on either characteristics of the claim (prescription drug spending, hospital care, GP care, care covered by supplemental health insurance) or characteristics of the submitting provider (laboratory tests). In order to separate a shared savings effect into a price and a volume effect, we reconstructed spending under a scenario of no price-differentials across providers. We did so by replacing each claim's price with the median price for that particular service [7].

In order to account for differences in characteristics of the insured between both groups, we created an aggregate measure of case mix which we denote as the 'risk score' (Appendix C), and controlled for this in our regressions by constructing risk score deciles.

3.3. Quality of care

We measured patient satisfaction at baseline (N = 309, 20 % response rate) and after the end of the first performance year (N = 595, 39 % response rate). The survey was send to all patients who were enrolled with a participating GP, and who had visited the GP in the three months prior to the start of the performance year or in the three months after the end of the first performance year (September, October, November 2015). We also evaluated the quality of chronic care delivery (diabetes and COPD) [12] and GPs'

prescription policy [13], using data from the GPs' registration systems.

3.4. Methods

For both total spending and categories of spending we used linear regression models with a difference-in-differences design [14–16] to isolate the effect of the shared savings pilot on quarterly spending (QS). We included patient fixed effects to account for pre-intervention differences in health care spending among patients (i.e. different 'risk profiles'). We included quarter fixed effects to account for time trends common to both groups. Because health care expenditures might develop differently over time for different risk deciles, the presence of which could violate the parallel trend assumption in case risk profiles are not equally distributed among GPs, which is highly likely, we also included linear time trends specific to each risk decile (modeled as deviations from the quarter fixed effects). Finally, We used GP-clustered, heteroscedasticity-robust standard errors [17], to account for correlation in expenditures between patients of the same GP, and serial correlation for the GP cluster [15,18]. Our estimation equation is:

$$QS_{igq} = \theta_i + \gamma_g + \tau_q + SSM_{gq}\beta + r_\rho q + \varepsilon_{igq} \quad (1)$$

In which QS refers to "quarterly spending", for individual i assigned to GP 'g' in quarter q. θ_i are patient-fixed effects, γ_g are GP fixed effects, and τ_q are quarter-fixed effects. 'SSM' is an interaction term consisting of an indicator variable of whether a patient's GP 'g' is part of the intervention group, multiplied by an indicator variable of whether quarter q belongs to the post-intervention period (which then functions as the intervention indicator). r_ρ are risk-decile specific deviations from the overall time trend τ_q (modeled linearly, note that q is a continuous variable denoting 'quarter'). Finally, ε refers to the error term of the equation.

The identifying assumption underlying a difference-in-differences design, is that pre-intervention spending trends would have continued in the absence of the pilot [15]. Although these numbers are counterfactual and cannot be observed, we tested the plausibility of this assumption by performing a falsification test [19]. Under such a test, one tests for deviations of a common trend in the pre-intervention period by adding leads and lags of the shared savings program. A significant lead effect implies a deviation from the common trend. In case one does not find a significant lead effect in the years before an intervention, and only observes deviating trends (right) after the onset of an intervention, the parallel trend assumption might hold. For this placebo test, our estimation equation is:

$$QS_{igq} = \theta_i + \gamma_g + \tau_q + AEZ_g\beta_{-9} + AEZ_g\beta_{-8} \ldots AEZ_g\beta_{start} + \ldots AEZ_g\beta_{+3} + r_\rho q + \varepsilon_{igq} \quad (2)$$

In which AEZ is an indicator variable of whether 'GP' was part of our intervention group, and in which the β subscripts denote the time to/after the start of our intervention (-10 is the reference period, which corresponds to the first quarter of 2012).

Our main specification reports the average effect of the shared savings program on spending, over all post-intervention periods. In order to assess whether this average accurately captures spending dynamics over time, we ran a difference-in-differences specification in which we interacted participation in the shared savings program with the time since the start of the program. As a sensitivity analysis, we ran a difference-in-differences model for each GP separately, to check whether a shared savings effect could be found for all GPs or only for a number of them. This analysis informs us about the generalizability of our findings.

Table 1
Summary statistics.

	AEZ	*Control group*
Number of insured	9,690	15,370
Age, yrs (+ SD)	44,92 (23)	43,49 (22)
Female sex (%)	51 %	51 %
Chronic illness in 2011 (%)		
Asthma	3,3%	2,9%
Diabetes Mellitus 2	2,9%	3,2%
Cardiovascular diseases	2,5%	2,6%
Risk score in 2011		
Mean	1,01	0,99
Interquartile range (Q3- Q1)	0,65	0,63
Average Quarterly expenditures (€)	668,83	649,70
Average Quarterly expenditures (capped) (€)	406,68	385,94

Table 2
Effect shared savings arrangement on quarterly spending, in Euro (€)*.

	Main specification	*.. under median prices*
Shared savings effect	−9.197**	−8.426*
(SE)	(3.745)	(4.264)
Observations	350,840	350,840
R^2	0.1623	0.171

NOTES: Standard errors in parentheses. Expenditures capped at 95th percentile.
* p < 0.10.
** p < 0.05.

Because control group data on quality were not available to us, we analyzed quality using a before-after design instead of a difference-in-differences design, using data from the treatment group only. Our quality indicators (Appendix B) cover patient satisfaction, chronic care delivery, and prescription drug policy. For each period, we calculate a weighted average per quality indicator, using the pre-intervention size of each participating GP as weights. This ensures that a comparison of quality before and after the first year of our program, is not confounded by changes in GP size. After all, each GP may have a different starting point.

4. Results

4.1. Summary statistics

Table 1 reports population characteristics separately for AEZ and the control group.

4.2. Descriptive evidence

Fig. 1 shows the time pattern of spending for both AEZ and the control group. On average, spending was €23 higher for AEZ (P < 0.001) in the pre-intervention period, and €15 higher (P < 0.001) in the post-intervention period. This difference of -€7 is not statistically significant (P = 0.11).

4.3. Change in medical spending

Despite this difference in absolute health care expenditures, we find that spending trends were parallel in the pre-intervention period: as shown in Fig. A1, the confidence intervals of the lead effects in the pre-intervention period always include the zero. This result supports a causal interpretation of our findings, as it shows that only after the onset of the pilot, spending trends changed.

Table 2 reports the results of our main specification. The sign and size of the difference-in-differences coefficient is negative and significant, suggesting that the shared savings program has led to

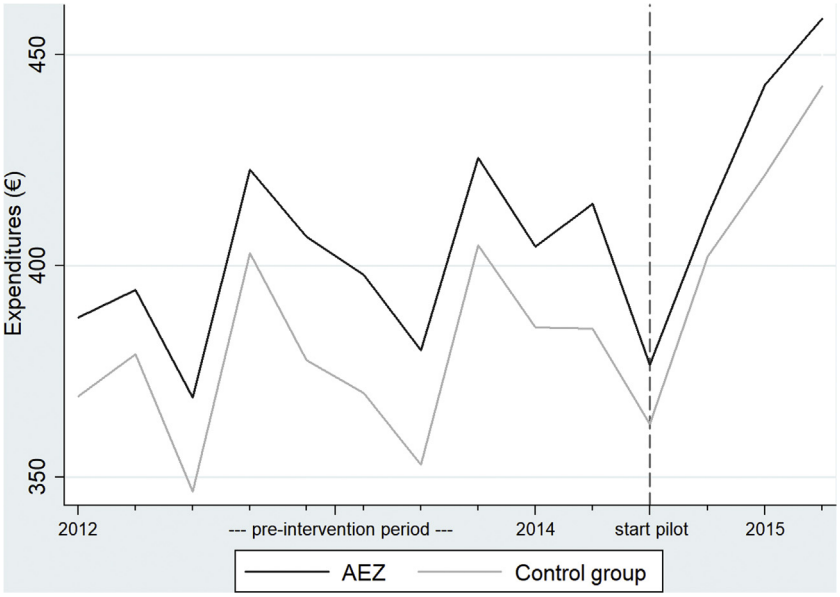

Fig. 1. Quarterly unadjusted medical spending for AEZ and the control group (capped at the 95th percentile).

Table 3
Effect shared savings arrangement on quarterly spending, according to type of care, in Euro (€).

	main specification
GP care	−0.54 (0.45)
Hospital care	−6.47*** (2.12)
Laboratory tests	−0.66** (0.24)
Spending under supplemental coverage	−0.21 (0.56)
Prescription drug spending	0.38 (0.55)

NOTES: Standard errors in parentheses. Expenditures capped at the 95th percentile.
** p < 0.05.
*** p < 0.01.

savings in health care expenditures. On average, and over the entire post-intervention period, savings equal €9 per person per quarter (P < 0.01). This corresponds to a 2 % reduction in average health care spending.

Fig. A2 plots the shared savings effect for each quarter separately. The figure shows that the effect is mainly concentrated in the first two quarters of the performance year. Although the coefficients in later quarters are of the right sign, their confidence intervals show that they are no longer statistically significant.

Fig. A3 shows that the expenditure savings were not concentrated within a single GP, but nearly every participating GP has displayed cost-saving behavior.

In order to disentangle price and volume effects, we imputed median prices for all services delivered to the population such that price differentials between providers vanish by construction. Our results are in Table 2, column 2. We find that the shared savings coefficient is still negative and significant. In terms of its size, the result suggest that savings have been predominantly achieved by lowering volumes.

Table 3 shows the results of additional difference-in-differences regressions, in which the dependent variable 'quarterly total spending', was replaced by category-specific spending only. We find a significant decrease in expenditures on hospital care and laboratory tests.

4.4. Change in quality of care (Appendices B/D)

We quantified changes in quality of care using a before-after study. We find small and statistically insignificant changes in patient satisfaction (p-value > 0.05), suggesting that patient satisfaction remained constant throughout our period of study. For chronic care and prescription drug policy, our results are ambiguous. We find both improvements and declines in quality. We find a large absolute decline in the percentage of patients with type-2 diabetes mellitus, who get treatment via a chronic care program (Appendix D).

5. Discussion

The results suggest that the shared savings arrangement was successful in changing GP behavior – which answers our first research question: the first performance year of the shared savings arrangement resulted in a 2% drop in per capita medical spending. Importantly, we find 5 out of 7 GPs displayed cost-saving behavior (for the other 2 GPs we find an increase in expenditures, but this increase is not significantly different from zero). This supports a causal interpretation of our findings.

We did however find significant drops or increases for some individual quality indicators. In discussing these findings with the participating GPs, we found no adequate explanation for these sudden swings in quality, suggesting that the shared savings arrangement has primarily targeted spending behavior and not quality. What is striking, however, is the large absolute decline in the number of patients enrolled in a chronic care program for diabetes (Appendix D). GPs enroll patients in to these programs, and receive a bundled payment for each enrolled patient, in exchange for which they have to provide chronic care. The GPs indicated that they critically reviewed their enrollment policy during the pilot period, in terms of patient no shows and the extent to which each diabetic patient would require the full service level as covered by the bundled payment. Patients with high no show rates, or patients who would be equally well off with a lower number of services (e.g. less foot exams), were excluded from the program and received regular care instead. This leads to direct cost savings, while it should have little or no effect on the quality of care. In its current form, the shared savings arrangement incentivizes to exclude no-show patients from chronic care programs (even when this implies a lower GP income), because these patients lower a GP's average score on the process indicators and directly lower the shared savings payment rate (which is a function of process qual-

ity). While this might seem counterintuitive at first glance, note that an increase in the savings rate – caused by higher quality – may well outweigh the loss in direct outcome.

As discussed earlier, one of the main motivations for implementing a shared savings arrangement, was to incentivize GPs to take up their role as coordinators within the health care system, by not only rewarding them for savings in primary care expenditures, but also for savings in spending by other providers. We find that the bulk of the shared savings effect can be attributed to a drop in hospital expenditures, suggesting that GPs have truly taken up this role. The large volume effect we find signals that GPs have realized a reduction in volumes, rather than merely shifting patients to lower priced care providers, as a response to the shared savings arrangement. Since the effect on quality remains ambiguous, it is impossible to say whether this is a drop in wasteful or low-value care only. A clear example of a reduction in wasteful care however, was the exclusion of no-show patients from chronic care programs. These results suggest that a (one-sided) shared savings arrangement is a promising way of transmitting value-based incentives to GPs, which answers our second and last research question. However, there are two issues that warrant further attention.

First, it is not yet clear whether the incentives that follow from a shared savings arrangement have a lasting impact on participating GPs. The dynamics of the shared savings effect showed that the arrangement's contribution to lowering health care spending was concentrated in the first two quarters of the year. The GPs themselves indicated that the most important intervention during the first performance year, was the sharing of expenditure data by Menzis. Although GPs are gatekeepers to specialist care, they are typically uninformed about medical spending elsewhere in the spectrum. Therefore, providing data on prices and volumes of spending outside the GP practices, and benchmarking these against other practices, was considered insightful by all GPs. This, however, was a one-off event because the possibilities for providing real-time data on expenditures were limited. It therefore seems likely that the effect of this intervention on GP behavior fades away with time, which is what we seem to observe. Whether sharing expenditure data prior to the start of the next performance year, or sharing data multiple times throughout the year, creates a similar or persisting effect on health care expenditure respectively, is a question open to follow-up research. Similarly, it is interesting to investigate whether data sharing is a mediator or a moderator of a shared savings effect, and under what conditions.

Second and last, we found that the impact of the shared savings arrangement on quality was ambiguous. Patient satisfaction was unaffected, while performance on other quality indicators was mixed – some increased, some decreased and some remained constant. Under the arrangement, a primary incentive for participating GPs is to lower spending, and, by letting the sharing rate depend on elements of performance, one can provide some constraints as to the interventions GPs will employ to earn shared savings payments. Nevertheless, the incentive to cut back on spending is stronger by design, because high quality, in and by itself, does not directly lead to shared savings payments. Depending on the situation at hand, and the nature of the value-based incentives that one wants to introduce via a shared savings arrangement, one may opt for a different way of intertwining costs and quality incentives. In our pilot group, quality levels were generally satisfactory while cost levels were not, so much weight was put on directing efforts to lowering spending, while, at a minimum, maintaining the quality status quo. This was done by giving large weights to 'absolute performance', relative to 'improvement'. However, the arrangement included a provision that a decline in quality of more than 5% would imply that no savings were shared at all. Stronger quality incentives can be introduced by putting more weight on improvement, or by sharing savings only conditionally on achieving a combined set of quality targets.

5.1. Limitations

This study has several limitations. A first limitation of this study is that the participating GPs were not selected at random, as they were all associated with a national organization of 'Arts en Zorg'. The idiosyncratic features of this organization, such as capacities related to data analysis, might not be generalizable to other practices. A difference-in-differences effect must therefore be interpreted as an average treatment on the treated, instead of an average treatment effect. A second limitation is that we were limited to performing a before-after study to detect changes in quality induced by the shared savings arrangement. Therefore, we are unable to separate a time trend – if any – from an intervention effect (using a difference-in-difference design, we would have been able to do so). In principle, some prescription drug spending indicators could be obtained from the claims data, but we lacked data on the prescriber, making it impossible to distinguish between instances in which a GP or a medical specialist (or other provider) made the prescription. A third limitation is that we capped expenditures at the 95th percentile of the spending distribution for each quarter, implying that, in principle, our results do not generalize to the full population of patients. In our sample, medical spending was skewed to the right. Spending outliers can have a profound influence on the spending average, which is the main economic quantity we are studying. In our case, outliers consumed up to 150 times more than the median and the top 5 percent spenders were responsible for 40 % of medical spending. Hence, even a slight imbalance in the number of outliers in the intervention and control group, can mask a shared savings effect. Surprisingly, the parallel trend test did not reject the null of parallel trends when using the uncapped expenditures. However, a closer look at the data revealed that, throughout our study period, the proportion of outliers was higher in the intervention group. Plots of the proportions of outliers in both groups showed that these differences were becoming smaller over time – a trend that was already visible at the start of our observation window (2 years before the start of the program). Using uncapped expenditures – which would be a legitimate choice given the results of the parallel trend test – would therefore overestimate the shared savings effect. Inasmuch the spending of outliers was limited in anticipation of joining the shared savings program, our results, using capped expenditures, are a conservative estimate of the true shared savings effect. A final limitation concerns the fact that using median prices might not do justice to the underlying differences in services with the same billing code.

6. Conclusion

This first year evaluation of the shared savings arrangement contains several insights for health care policy. First, a shared savings arrangement can help transmitting value-based incentives to GPs, as it caused GPs to bend the cost curve while patient satisfaction remained at equal levels. Second, performance under a shared savings arrangement seems to be driven by the extent to which it is possible to share spending data with GPs, suggesting an important role for privacy law here. Third, even in a highly fragmented health care system, one may be able to find provider types who are able to assume accountability for total health care spending. Thus, market competition does not necessarily have to be compromised, in the sense that one does not need vertical integration of providers across the spectrum in order to assume accountability [20]. Fourth, we also found a drop in GPs' own expenditures, suggesting that it is not necessarily true that the fee-for-service incentive is stronger

than the shared savings incentive [21]. We attributed this drop in GP expenditures in part to the process indicators in our pay-for-performance element in the arrangement, which discourage GPs from billing a bundled payment for patients that do not require or do not want the full bundle of services. Here, the GPs might have traded off a loss in revenue from billing fewer bundled payments against a higher sharing rate resulting from higher reported quality. Finally, this evaluation also confirms that a one-sided risk arrangement, in which GPs do not share in excess spending, already elicits substantive effort to reduce spending – even when it is introduced as a 'mere' complement to the existing payment model [8].

Declaration of Competing Interest

AH finished this study while working for Menzis. Academic objectivity was guaranteed by means of a separate contract.

Acknowledgements

We acknowledge funding from Stichting Menzis/Azivo RVVZ, the Netherlands.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.healthpol.2021.01.013.

References

[1] Miller HD. From volume to value: better ways to pay for health care. Health Affairs 2009;28(5):1418–28.
[2] Zuvekas SH, Cohen JW. Fee-for-service, while much maligned, remains the dominant payment method for physician visits. Health Affairs 2016;35(3):411–4.
[3] Kroneman M, Boerma W, van den Berg M, Groenewegen P, de Jong J, van Ginneken E. Netherlands: health system review. Health Systems in Transition 2016;18(2):1–240.
[4] Berwick DM. Launching accountable care organizations—the proposed rule for the Medicare Shared Savings Program. New England Journal of Medicine 2011;364(16):e32.
[5] Chernew ME, Mechanic RE, Landon BE, Safran DG. Private-payer innovation in Massachusetts: the 'alternative quality contract'. Health Affairs 2011;30(1):51–61.
[6] Ginsburg PB. Spending to save—ACOs and the Medicare shared savings program. New England Journal of Medicine 2011;364(22):2085–6.
[7] Song Z, Safran DG, Landon BE, Landrum MB, He Y, Mechanic RE, et al. The 'Alternative Quality Contract,'based on a global budget, lowered medical spending and improved quality. Health Affairs 2012;31(8):1885–94.
[8] McWIlliams JM, Hatfield LA, Chernew ME, Landon BE, Schwartz AL. Early performance of accountable care organizations in Medicare. New England Journal of Medicine 2016;374(24):2357–66.
[9] Hayen AP, van den Berg MJ. Incorporating shared savings programs into primary care: from theory to practice. BMC Health Services Research 2015;15(1):580.
[10] Centers for Medicare & Medicaid Services. Medicare program; medicare shared savings program: accountable care organizations. Final rule. Federal Register 2011;76(212):67802.
[11] Conrad DA, Perry L. Quality-based financial incentives in health care: can we improve quality by paying for it? Annual Review of Public Health 2009;30:357–71.
[12] Ineen. Transparante ketenzorg; 2014.
[13] Instituut voor Verantwoord Medicijngebruik. Benchmark verantwoord voorschrijven; 2014.
[14] Card D, Krueger AB. Minimum wages and employment: a case study of the fast food industry in New Jersey and Pennsylvania. National Bureau of Economic Research; 1993.
[15] Angrist JD, Pischke J-S. Mostly harmless econometrics: an empiricist's companion. Princeton university press; 2008.
[16] Caliendo M, Hujer R. 14 The microeconometric estimation of treatment effects-an overview. Econometric Analysis 2006:199.
[17] White H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica: Journal of the Econometric Society 1980:817–38.
[18] Bertrand M, Duflo E, Mullainathan S. How much should we trust differences-in-differences estimates? The Quarterly Journal of Economics 2004;119(1):249–75.
[19] Autor D, Katz L, Kearney M. Rising wage inequality: the role of quantities and prices. NBER working paper; 2005. p. 11628.
[20] McWIlliams JM, Hatfield LA, Chernew ME, Landon BE, Schwartz AL. Performance differences in year 1 of pioneer accountable care organizations. New England Journal of Medicine 2015;372(20):1927–36.
[21] Berenson RA. Shared savings program for accountable care organizations: a bridge to nowhere? The American Journal of Managed Care 2010;16(10):721.
[22] Kringos DS, Boerma W, van der Zee J, Groenewegen P. Europe's strong primary care systems are linked to better population health but also to higher health spending. Health Affairs 2013;32(4):686–94.
[23] Nederlandse Zorgautoriteit. Marktscan Huisartsenzorg; 2012.
[24] Schellevis FG, Westert GP, De Bakker DH. The actual role of general practice in the Dutch health-care system. Journal of Public Health 2005;13(5):265–9.
[25] Schäfer W, Kroneman M, Boerma W, van den Berg M, Westert G, Devillé W. Health systems in transition; 2010.
[26] Schut FT, Varkevisser M. Competition policy for health care provision in the Netherlands. Health Policy 2017;121(2):126–33.
[27] McWilliams JM, Schwartz AL. Focusing on high-cost patients—the key to addressing high costs? The New England Journal of Medicine 2017;376(9):807–9.
[28] Centers for Medicare and Medicaid Services, et al. The physician group practice transition demonstration Bonus methodology specifications. RTI International; 2004.
[29] Erez M, Zidon I. Effect of goal acceptance on the relationship of goal difficulty to performance. Journal of Applied Psychology 1984;69(1):69.
[30] Pope GC, Kautter J. Minimum savings requirements in shared savings provider payment. Health Economics 2012;21(11):1336–47.