Abstract

Paying on the basis of fee-for-service is often associated with a risk of overprovision. Policymakers are therefore increasingly looking to other payment schemes to ensure a more efficient delivery of health care. This study tests whether context plays a role for overprovision under fee-for-service. Using a laboratory experiment involving medical students, we test the extent of overprovision under fee-for-service when the subjects face different fee sizes, patient types, and market conditions. We observe that decreasing the fee size has limited effect on overprovision when resources are abundant, but it does significantly affect provision when resources are constrained. We also observe that patients who are harmed by excess treatment are at little risk of overprovision. Finally, when subjects face resource constraints, but still have an incentive to overprovide high-profit services, they hesitate to do so, implying that the presence of opportunity costs in terms of reduced benefits to other patients protects against overprovision. Thus, this study provides evidence that the risk of overprovision under fee-for-service depends on fee sizes, patients’ health profiles, and market conditions.

Keywords: Fee-for-service, supplier-induced demand, laboratory experiment
JEL-codes: C91, I11

Acknowledgements: Financial support from the FAIRCARE research programme funded by University of Southern Denmark and the Danish Council for Independent Research | Social Sciences (Grant ID: DFF-6109-000132) is gratefully acknowledged. We would also like to thank participants and discussants at the 39th Nordic Health Economists’ Study Group Meeting, the 13th Nordic Conference in Behavioural and Experimental Economics, and the Value of Health Conference for helpful comments.
1. Introduction

In many health care systems physicians are paid a fee per service (FFS) they provide. FFS is popular because it gives physicians a direct financial incentive to exert effort into health care. However, FFS may lead to overprovision of care making it difficult for policymakers to contain costs (Dranove 1988). Consequently, alternative payment forms, such as pay-for-performance, are gaining popularity (WHO, 2014).

Several studies confirm the existence of overprovision under FFS (e.g. McGuire 2000; Madden 2005; van Dijk et al. 2013; Hennig-Schmidt et al. 2011; Brosig-Koch et al. 2017). These findings may, however, not be robust across patients, fee sizes, and market conditions. Our study examines whether context matters for overprovision to arise under FFS – and thus whether its somewhat poor reputation is always justified. We hypothesize that there exist so-called protective factors against overprovision.

Physicians’ utility is typically assumed to depend on both their own net revenue and patients’ health (Arrow 1963; Ellis & McGuire 1986). In cases where overprovision of care is directly or indirectly detrimental for patients’ health, FFS forces physicians to trade-off patients’ health for net revenue. Given physicians’ altruism (Ellis & McGuire 1986; Godager et al. 2013), we test whether patients’ health loss may serve as a protective factor against overprovision under FFS.

Under FFS, the fee is typically higher than the marginal costs of care, implying that physicians’ gain a profit from providing services. When there is a trade-off between patients’ health and physician’s profit a lower fee entails that physicians sacrifice less to prioritise patients’ benefit over own profit. Consequently, we test whether lower fees per service provided lead to a reduced overprovision of care.

Overprovision has mainly been observed in settings in which physicians are resource abundant, i.e. they have the equipment and personnel required to exhaust their patients’ need of care (e.g. Hennig-Schmidt et al. 2011; Brosig-Koch et al. 2017). However, many physicians experience resource constraints on a daily basis, such that care to one patient comes at the expense of others. Studies have investigated physicians’ allocation of resources under constraints, but without testing the existence of overprovision under FFS (e.g. Ahlert et al. 2013; Ahlert and Schwettmann 2017; Damm et al. 2014; Hurst et al. 2005). We fill this gap by testing whether overprovision is reduced when resources are constrained compared to a situation of resource abundance under FFS.

An increasing number of studies on the provision of health care make use of fully-incentivised laboratory experiments with medical students (Galizzi and Wiesen 2017). These studies have confirmed findings from natural experiments and expanded the analyses to include additional features, e.g. in relation to payment schemes (Hennig-Schmidt et al. 2011; Brosig-Koch et al. 2017). The advantage of using the laboratory setting is that it allows for testing the effect of a change in a single feature, which may be difficult in a natural setting where many features may be fixed, unobservable, or change simultaneously. In this study we therefore also make use of a laboratory experiment to study the importance of context for overprovision under FFS. Section 2 presents our experimental setting, Section 3 our main findings, and Section 4 concludes.

---

1 E.g. physical and psychological stress.
2 Natural experiments depart from laboratory experiments in that participants act in their natural environment without knowing that they take part in an experiment (Galizzi and Wiesen 2017).
2. Experimental setting

2.1 Experimental procedure

We ran an incentivised computer-based experiment with 38 medical students (42% male, average age 23)\(^3\). The experiment was run at SamExpri, the experimental laboratory at the Faculty of Social Sciences, University of Southern Denmark. An ex-ante power calculation showed that 37.8 subjects were sufficient to guarantee a power of 0.85.\(^4\) Participation was voluntary, and subjects were paid on average DKK 126 (USD 20), which included a show up fee of DKK 30 (USD 5). As the experimental framework referred to an interaction between physician and patients, all services that the subjects provided to the fictitious patients were transformed into money that were subsequently donated to two charitable health-related organisations.\(^5\)

In the experiment, subjects played the role of physicians, and were asked to allocate services among patients. Physicians worked for 36 “days”, on which they consulted either one or two patients. Each “day” physicians could allocate a maximum of 12 units of services in the way they considered best. In the experiment, physicians incurred no costs of treating patients and were paid based on the number of services they were providing. The payment should thus be interpreted as the net revenue generated per service provided.\(^6\) Once the maximum capacity to benefit was reached, additional services had either no effect on health or were detrimental to health.

Both at the start and at the end of the experiment, participants filled in a questionnaire proving, among other things, whether their task was understood.

2.2 Definition of patient types

Patients were defined by two characteristics: the minimum number of services needed to exhaust their capacity to benefit (CTB) and whether they were harmed in case of overprovision. We defined overprovision as care exceeding the minimum amount required to exhaust the patient’s CTB. Patients could either be characterised by high-CTB (high responsiveness to treatment, requiring 5 services to reach maximum benefit) or low-CTB (low responsiveness to treatment, requiring 10 services). The maximum CTB was the same for all types of patients. However, once the maximum CTB was reached, patients could either be unharmed or harmed by overprovision. We defined “harm” as a decrease in the benefit from care when receiving an additional service. This design allowed us to capture how physicians act when facing patients with different health profiles. The benefits from receiving services for each of the four patient types are illustrated in Table 1.

As we aimed to test how the fee size impacts provision of care, each patient type was associated with two different fees: either $1 (experimental $) per service or $2 per service.

---

\(^3\) 40 students were recruited. One student left before signing the informed consent, while data from another were removed from the dataset after the participant self-reported to have incorrectly entered the values in the software.

\(^4\) For more details on the power calculation see appendix A1.

\(^5\) For more details on the experimental protocol and design see appendices A3 and A4.

\(^6\) In principle we could have designed the experiment such that each service triggered both a cost and a payment – and let the subject calculate the profit margin, but to minimize the cognitive burden on subjects, we chose this simpler and equivalent design.
Physicians faced every possible combination of patients (individually and in pairs), for every possible combination of payment. This resulted in a total of 36 working “days”; 8 “days” consulting only one patient and 28 “days” consulting two patients. The two-patient “days” illustrated a situation of resource constraints (except for combinations of two high-CTB patients) where physicians were forced to trade-off care between patients.7

**Table 1:** Patient’s health gain for a given number of health care services

<table>
<thead>
<tr>
<th>Number of health care services</th>
<th>Unharmed Low-CTB</th>
<th>Unharmed High-CTB</th>
<th>Harmed Low-CTB</th>
<th>Harmed High-CTB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>8</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>10</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>10</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>10</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>3</td>
</tr>
</tbody>
</table>

3. Results

Initially, we analysed the effect of fee size and patient type in the case when physicians were resource abundant. Subsequently, we measured the overall effect of introducing a resource constraint, and in particular analysed whether the role of fee size and patient type changed under this new market condition. The following sections present our results by market condition.8

3.1 Resource abundant physicians

Around one-third of the patients were overprovided when physicians were resource abundant (mean=35%; CI=[25%; 46%])9; however, the magnitude of the overprovision was heterogeneous across patient types. Unharmed patients were overprovided on average 57% of the time, against a 13% of harmed patients (p<0.001, repeated measures logistic regression). Furthermore, the number of services overprovided was on average 2 units higher for the unharmed patients than for the harmed ones (p<0.001, two-sided t-test10), see Figure 1. Thus, in this setting patients’ health acted as a protective factor against overprovision.

Reducing the fee size was not a strong protective factor against overprovision. Both the harmed, low-CTB patients and the harmed, high-CTB patients experienced no statistically significant change in the frequency of overprovision when the fee size was reduced. The average number of overprovided services for these harmed patients was less than one unit, see Figure 1.

---

7 For completeness our analyses of “resource constraints” were based on all combinations of patient types. However, our findings were robust to the exclusion of “days” with two high-CTB patients.
8 For more details on our empirical approach and results see appendices A1 and A2.
9 All proportion confidence intervals are Clopper-Pearson binomial CI-intervals with standard errors clustered at physician-level.
10 Standard errors obtained by bootstrapping (resampling clustered by physician, 10,000 replications).
3.2 Resource constrained physicians

When physicians were resource constrained we observed both overprovision and underprovision. First, we analysed the case of overprovision. On average, only 8% (CI=[6%; 11%]) of patients were overprovided when physicians were resource constrained. The overprovided patients were primarily characterised by being the unharmed, high-CTB patients (mean=24%, CI=[19%; 31%]). However, the harmed, high-CTB patients were also slightly overprovided (mean=5%; CI=[2%; 10%]). On average, unharmed patients were 10 percentage points more at risk of being overprovided than harmed patients (p<0.001, repeated measures logistic regression). This finding indicates that patients’ health serves as a protective factor against overprovision also when physicians’ resources are limited.

We observed that a lower fee significantly reduced overprovision for the unharmed, high-CTB patients, reducing the frequency of overprovision by 14 percentage points (p<0.001, repeated measures logistic regression). A similar trend was observed for the harmed, high-CTB patients, where overprovision was eliminated when the fee was reduced (frequency of overprovision reduced by 6 percentage points, p<0.001). Thus, when physicians are resource constrained, the fee size may protect patients who are at risk of overprovision, i.e. the high-CTB patients.

The natural downside to physicians being resource constrained is that some patients are underprovided. Figure 1 illustrates that it is especially the low-CTB patients – both harmed and unharmed – that were underprovided. More specifically, we found that nearly all low-CTB patients were underprovided (mean=97%, CI=[94%; 100%]). However, the risk for low-CTB patients decreased by 5 percentage points when physicians received a high fee versus a low fee (p<0.001, repeated measures logistic regression). We therefore conclude that increasing the fee may serve as a protective factor against underprovision of care to low-CTB patients under resource constraint.

**Figure 1:** Average number of services provided to each patient type for different fee sizes under either resource abundance or resource constraint

Note: "Resource abundant" refers to the one-patient scenario, where physicians were able to exhaust patients’ need of care. "Resource constrained" refers to the two-patients scenario, where physicians were forced to trade-off care between patients (except for pair of high-CTB patients). "High" ("Low"): average number of services provided to patients when physicians were paid a high (low) fee. "Optimal care": minimum number of services that maximises the patient’s health benefit from care.
4. Conclusion and discussion

This laboratory experiment illustrated how context influence the extent of overprovision under FFS. Needless to say, a laboratory setting cannot mimic the real world, and thus our results should be interpreted in qualitative terms. Previous laboratory experiments have successfully confirmed and expanded findings from natural experiments (e.g., Hennig-Schmidt et al. 2011, Brosig-Koch et al. 2017). Our study confirmed the previous finding of overprovision under FFS. However, it also showed that under resource constraint overprovision was less frequent than under resource abundance. One may argue that this finding is by construction, however, the incentive to provide services to patients who generate a higher profit was still in effect under resource constraint. Thus, our results imply that “patient opportunity costs” offer protection against overprovision. Moreover, while the fee size seemed to have limited effect on the risk of overprovision under resource abundance, it may be used as a tool to both reduce the risk of underprovision for low-CTB patient and overprovision for high-CTB patients under resource constraints. Finally, patients who were harmed by overprovision were less at risk of being overprovided, suggesting that the health risk embedded in FFS could be more limited than previously thought. We conclude that because of these structural protective factors against both under- and overprovision, FFS appears better than its reputation in many contexts. Thus, when deciding whether or not to make use of FFS, both the characteristics of the patient population, the fee size, as well as resource availability should be considered.

5. References


