State Dependence in Health Care Provider Choice: Evidence from Dental Care

Tuomas Markkula*

November 16, 2024

Click here for the most recent version

Abstract

Health care markets are famously imperfect. One of the reasons is the possibility that consumers are facing choice frictions like switching costs, brand loyalty, or habit formation. Choice frictions lead to true state dependence in consumer's choices, a pattern in which the previously visited provider becomes more desirable today because it was visited previously. I test for the existence of the true state dependence in health care provider choice in the context of Finnish private dental care. Switching is rare in this market, which can be due to the true state dependence, or because of consumers' persistent preference heterogeneity for dental practices. I disentangle the two competing explanations by controlling for consumers' unobserved preferences with consumer-dental practice fixed effect and find evidence of true state dependence with a magnitude similar to a three-and-a-half-kilometer reduction in distance to the previously visited dental practice. This is a sizable effect considering that the average distance from a consumer to the practice they are visiting is 16 kilometers.

*I am grateful to Otto Toivanen, Tanja Saxell, and Ciprian Dominisoru for supervision, and to Janne Tukiainen, Jiekai Zhang, Mika Kortelainen, Markku Siikanen, and Nea Leinonen for useful comments. I thank HSE Support Foundation, Yrjö Jahnsson Foundation and OP Research Foundation for generous funding.

1 Introduction

Health care markets are notoriously imperfect, primarily due to asymmetric information that results in issues such as moral hazard, adverse selection, and supplier-induced demand (Arrow, 1963). Another possible source of market imperfections is choice frictions, which deter consumers from switching providers despite potential benefits. These choice frictions take many forms, including switching costs, search costs, and inattention. In health care, beneficial relationships between a health care provider and the patient is a possible source for the choice frictions, as demonstrated by Sabety (2022). Reluctance to switch providers can lead consumers to make suboptimal choices, influence the nature of market competition, and impact policymaking, particularly concerning mergers and antitrust (Dubé et al., 2009; Handel, 2013; MacKay and Remer, 2022). Finally, the presence of choice frictions can be a source of market power for firms (Farrell and Klemperer, 2007). Yet, the importance of choice frictions in primary health care is relatively unknown.

I test for the existence of the choice frictions[1] in primary health care provider choice, by evaluating whether consumers' choices exhibit true state dependence in the context of Finnish private dental care. The true state dependence in choices means that after visiting a dental practice, the probability of choosing the practice in the next period increases (Heckman, 1981). The existence of choice frictions results in consumer choices that exhibit true state dependence. The main empirical challenge is that even though I show that switching is rare in this market, consumers' persistent choices can result from consumers' persistent preference heterogeneity for dental practices in addition to the causal true state dependence. I disentangle the two competing explanations with Honoré and Kyriazidou (2000) approach that controls for consumers' unobserved preferences with consumer-dental practice fixed effect in a multinomial logit model. I find evidence of true state dependence with a magnitude similar to a three-and-a-half-kilometer reduction in distance to the previously visited dental practice. This is a sizable effect considering that the average distance from a consumer to the practice they are visiting is 16 kilometers.

The Finnish private dental care is an excellent setting for studying choice frictions in a health care context. First, in Finnish private dental care the dental practice choice does not depend on an insurance plan. This makes it simple to analyze practice choice separately from insurance plan choice. Second, maintaining dental health requires frequent visits to a dental practice. This allows me to observe several visits for each consumer in my data,

[1] I view the choice frictions as resulting from any combination of the following more specific ways in which the consumers' choices are affected by their previous choices: switching costs, brand loyalty, habit formation, search costs, learning, and inattention.

making it easy to determine which dental practice is the consumer's current practice and when they switch. Third, dental care is an important health care service, that is among the most used primary care services for non-elderly (Buchmueller et al., 2016). Thus, the possible existence of choice frictions in this health care service has implications for everyone's health care service use.

I use rich consumer-level panel data on visits to Finnish private dental care between 2008 and 2017 to document that most consumers return to their previous practice. Out of all consumers who have at least two treatment episodes in the data, 95% chooses their previous period's practice. Even though the pattern in the data is suggestive of the true state dependence, a similar choice pattern could also be produced by consumers with persistent unobserved preference heterogeneity towards practices (Heckman, 1981). This possible confounder of the true state dependence is called spurious state dependence. The distinction is important, as the true state dependence implies that firms can increase their future demand through discounts. A consumer who uses a discounted service is more likely to use the service again later when the discount has ended. On the other hand, the spurious state dependence implies that the consumer's purchase during a discount has no effect on their future choices, as choices are driven by time invariant preferences. Thus, whether the repeated choices are due to the true state dependence or due to the spurious state dependence is of substantive importance.

I disentangle the two competing explanations for consumers' persistent choices by controlling for consumers' unobserved preferences with consumer-practice fixed effects in a multinomial logit using the sufficient statistic approach of Honoré and Kyriazidou (2000). Consumer-practice fixed effects allow me to control for practice-specific, consumer-specific, and consumer-practice-specific time-invariant variables. Standard Berry et al. (1995) type demand models estimated using market-level data allow controlling for practice-specific time-invariant variables. Access to consumer-level data allows a researcher to control for consumer characteristics, like income, and interactions between practice and consumer characteristics, like price and income. However, controlling for consumer characteristics and adding interactions requires observing practice and consumer characteristics. In contrast, my approach controls for all consumer-specific time-invariant factors that are typically unobserved, like tastes. Furthermore, consumer-practice-specific fixed effects control for interactions between consumer's time-invariant tastes and all unobserved time-invariant characteristics of the practice, like waiting room decor.[2]

[2]Even though time varying consumer and practice characteristics are not controlled for by the consumer-practice fixed effects, it captures the general level of the variables that do not change dramatically from one year to another like consumer's income and age.

In addition to the fixed effects, I control for two important time-varying covariates affecting practice choice: practices' prices and consumers' distance to the practices. Including distance in the model allows me to express the magnitude of the true state dependence in terms of traveling distance. The fixed effects model imposes stringent requirements on the data, shrinking the sample from a panel of five and half a million treatment episode observations down to a cross-section of 4 615 sequences of choices with three or four observations each. Fortunately, the restricted and the full sample are similar in observable characteristics.

I find evidence of the true state dependence when using a model with consumer-practice fixed effects and controlling for distance. The magnitude of the state dependence is similar to a three-and-a-half kilometer reduction in distance to the previously visited clinic. This is a significant effect, considering that the average distance between a consumer and the practice they visit is 16 kilometers. When I introduce price to the model the coefficient increases moderately but is no longer statistically significant. This could be because of omitted variable bias of not including price into the model or from an observation weighting scheme that Honoré and Kyriazidou (2000) approach requires when adding variables into the model that change over time. I argue that the reason is the weighting as different weighting schemes only increase the point estimate of the true state dependence. Moreover, consumer-practice fixed effects should control any time-invariant source of price endogeneity including the correlation between unobserved quality and prices.

In light of my results and the previous literature finding similar results, Raval and Rosenbaum (2018) for choice of hospitals for childbirth and Irace (2018) for hospital choice, it seems likely that choice frictions are an important reason alongside imperfect competition why health care markets might fail. Policymakers should be aware of the existence of choice frictions in health care markets and take them into account. Finally, whether it might be worthwhile to try to reduce the choice frictions in health care markets through policy depends on the type of choice frictions. If choice frictions stem from the experience good property of health care then a quality rating system might help to reduce choice frictions. On the other hand, if choice frictions arise from improved quality of health care due to better continuity of care, reducing choice frictions might not be advised.

Related Literature I contribute to three strands of literature. First, I contribute to the literature on choice frictions in health care services by estimating the importance of true state dependence and choice frictions in a primary care setting, which has not been done before to my knowledge. Closest to my work are studies estimating switching costs for physicians. Kwok (2024) identifies the monetary switching costs of switching a physician for elderly Medicare patients from a change in health care utilization after a switch of physician.

While my estimate of choice frictions contains the direct monetary costs of changing a dental practice, it also includes other choice frictions. Moreover, I focus on adults instead of the elderly. Dahl and Forbes (2023) use a quasi-experiment that increased the cost of keeping a patient's current physician by $600 to $1900 to evaluate how much consumers are willing to pay to keep their current physician. They refer to their estimates as switching costs, but as they are not able to control for the patients' preferences towards their current doctor they are not able to separate consumers' persistent preference heterogeneity from true state dependence. I differ from Dahl and Forbes (2023) in that I estimate the true state dependence in a primary care context instead of a combination of preferences for a physician and switching costs. Finally, Anell et al. (2021) performed an experiment that reduced switching costs and other choice frictions in the context of primary care providers in Sweden and found that the experiment increased switching. However, they do not estimate the magnitude of switching costs, which I do.

Other literature on health care markets considers choice frictions in hospital choice (Raval and Rosenbaum, 2018; Irace, 2018), drug choice (Granlund, 2021; Feng, 2022; Janssen, 2022) and health insurance plan choice (Abaluck and Gruber, 2011; Handel, 2013; Polyakova, 2016; Yeo and Miller, 2018; Pakes et al., 2021; Heiss et al., 2021; Tilipman, 2022). Second, I contribute to the literature studying the dental care industry by demonstrating the existence of choice frictions in the industry. This literature includes works like Choi (2011), Decker and Lipton (2015), Wing and Marier (2014) and Buchmueller et al. (2016). Third, my work is related to the large literature studying choice frictions in various markets like (Erdem and Sun, 2001; Chintagunta et al., 2001; Dubé et al., 2010; Kong et al., 2022) for consumer packaged goods, Hortaçsu et al. (2017) for residential electricity providers, Grzybowski and Nicolle (2021) for smartphones and Luco (2019) for retirement investment.

2 Industry, Data and Consumers' Switching Behavior

The Finnish private dental care industry is a good setting to study state dependence in provider choice, as the insurance structure does not affect provider choice. Moreover, the Finnish visit level microdata allows me to observe essential consumer, practice, and dentist identifiers and possible confounders of true state dependence like whether consumer moves or whether firms change their prices. With this data, I show that switching a dental care practice is rare.

2.1 The Finnish Dental Care Industry

In the Finnish dental care industry, consumers can choose between publicly or privately provided dental care practices that share the market with equal market shares (Nurminen et al., 2021). Studying practice level switching in the private sector is facilitated by the fact that consumers can freely choose among private dental practices competing for customers with location, quality, and price.

The expenditures in the Finnish dental care markets were around 1 billion euros in 2019, which is roughly 4.5% of all health care expenditures in Finland during 2019 (The Finnish Institute for Health and Welfare, 2021a). The markets for dental care are local and as most consumers visit a practice in their municipality, I assume that every municipality is their own market. The only exception is the Capital Region where four municipalities[3] form one market.

Both private and public practices are present in most of the markets, and they tend to share the market with equal market shares. The two sectors produce essentially the same set of dental care services. Switching between the sectors is rare (Nurminen et al., 2021). Municipalities are mandated by law to provide public dental care. Larger municipalities tend to have several public clinics providing dental care services, while smaller municipalities often have only one clinic. Public dental care has low copayments. In 2015 the copayment for the basic dental examination was capped nationally at 14,70 euros. However, waiting times for public sector care can be long. In October 2015, out of all non-emergency care visits to public dental care in Finland, 36.6% had a waiting time over 21 days and 13.4% had a waiting time over three months (The Finnish Institute for Health and Welfare, 2021b,c). The waiting times vary across municipalities.

Alternatively, consumers can choose from several private firms. There are few large national chains operating larger clinics and many independent practitioners who practice dentistry by themselves or in small group practices with less than five dentists. In all but the smallest markets, there are many private firms consumers can choose from. In the biggest cities like Tampere, which is among the three largest urban areas in Finland, the number of private practices can be as large as 65, while in the capital Helsinki, there are over a hundred different private practices present. Private practices' pricing is not restricted, and private sector services are covered by the National Health Insurance, which covers all Finnish citizens and has only one plan that does not restrict practice choice. The insurance structure allows me to isolate the practice level true state dependence from the insurance plan level true state dependence.

[3]These municipalities are Helsinki, Espoo, Vantaa, and Kauniainen.

Relative to the public sector, waiting times in the private sector are short, but services are more expensive. For example, in 2015 the nationwide average nominal price of a dental examination was 62.60 euros. The National Health Insurance reimburses consumers' private sector expenses by a small, fixed amount that is treatment-specific. The insurance reimbursement for examination was 15.50 euros in 2015, and hence the consumer ended up paying 47.10 euros, which is three times the maximum copayment in the public sector during the same year. Private sector services also differ from public sector services in that consumers can choose which dentist they want to visit. Most of the time when a consumer returns to their previous practice, they also choose the same dentist as during the previous visit. As several mechanisms giving rise to the true state dependence are patient-dentist pair specific, this is a convenient feature of the market that allows me to capture in my estimate both true state dependence related to the practice and that related to the dentist, as long as the dentist does not move to another practice. In the public sector, consumers are not allowed to choose their dentist, and instead, the consumer is allocated to one of the dentists working in a chosen clinic. Consumers choose a practice to visit based on price, location, and perceived quality. Thus, it is not surprising that consumers who frequent private firms tend to have higher incomes than consumers frequenting public practices.

2.2 Consumer Level Panel on Visits to Private Dental Care

I use consumer-level panel data on the private sector's dental examinations in Finland between 2008 and 2017 that were reimbursed by the National Health Insurance to study true state dependence in practice choice. The data contains five and a half million dental examinations, which I use as a proxy for dental treatment episodes.

The data comes from the Social Insurance Institution of Finland Kela, and it contains rich information on consumers, dentists, and dental care practices that allow me to estimate practice switching costs reliably. The data is at the procedure level. The rows in the data contain the following information on the procedure: the procedure, the procedure's daily price at the practice, the size of the National Health Insurance reimbursement for the procedure and whether the procedure was performed by a specialized dentist, and whether the procedure was performed at an unusual time.[4] If the two latter conditions apply, the consumer is eligible for an additional National Health Insurance reimbursement. I observe all this information for the dental examinations. I convert the prices to real quantities with Statistics Finland's Consumer Price Index using 2005 as the base year. For consumers, I observe identifiers, age, sex, yearly earnings, capital income, and location of residence at

[4]The unusual times are during the week between 21.00–7.00, on weekends, or after 18:00 on the eve of national holidays or during a national holiday.

the zip-code level. Practice identity is defined as a combination of a unique practice name and a zip code, so I also observe practice location at the zip code level. For dentists, I observe identifiers, sex, age, graduation year, and medical specialties. A remarkable feature of the data is that for almost all dental examinations I observe who the patient is, who the treating dentist is, and at which practice the examination was done. Thus, I observe when a consumer switches practice, which is essential when studying state dependence on choices.

I use dental examinations as a proxy for a treatment episode because they contain sufficient information to study consumer's practice choice behavior, and because using only examinations allows me to avoid some unnecessary complications in the empirical analysis. My data contains all reimbursed dental care procedures performed at the private sector, but using only examinations is sufficient because almost always all other procedures are performed at the same practice where the examination was performed. This is because all information that the practice obtains during the examination is not transferable, and thus the new practice might require a new examination before any treatments. Moreover, focusing on examinations simplifies the analysis because I do not need to take into account the fact that consumers might differ in what treatments they need. Likely, consumers with worse dental care routines and more time since the last visit are in need of more extensive treatments, and thus of also more expensive treatments. Taking this into account would require modeling a consumer's dynamic decision whether to go to the dentist this year or not, which is not essential for estimating the extent of true state dependence in practice choice.

The National Health Insurance reimburses almost all dental examinations performed in private dental care. The most important restriction of the insurance is that for each consumer, it only reimburses one dental examination per year. This means that I do not observe if a consumer visits a dentist twice within a year. Fortunately, most of the consumers visit the dentist at most once a year as that is the visit frequency recommended by dental care literature, for example by Giannobile et al. (2013), and also by one of the Finnish chains, Mehiläinen (2022). Nevertheless, I might be missing some visits for individuals with a need for more frequent dental care examinations, but this is likely a small share of the total population. Additionally, from 2015 onward the insurance only reimburses a dental examination once in two years, which restricts how I can use the last years in my data: 2015, 2016, and 2017. I will not be using the year 2017 in the main analysis.

Another limiting feature of the data is that I am missing practice identity for some observations. I only observe the identity if the practice in question has taken up the direct reimbursement system for the National Health Insurance. I am not able to use an observation if it is missing practice identity, and hence I drop these observations. The problem is worse for the early years of the data, as the take-up of the direct reimbursement increases over time.

In 2008 20.7% of the examinations are missing practice names, but by 2017 the percentage drops to 5.7%.

2.3 Switching a Dental Practice is Rare

If consumers switch practices often, then that indicates that there might not be true state dependence or that the dependence is small. On the contrary, I find that switching is rare, as 95% of consumers choose their previous period's practice out of all consumers who have at least two treatment episodes in the data.

I observe many practice switches in the data, but the amount is small relative to the total number of observations in the data. I define practice switching to mean a choice such that a consumer visits a different practice during this visit than during the previous visit. Only consumers who have visited at least twice can switch, so I focus on these customers. There are five million visits in the data from consumers who have visited at least twice. Out of the five million visits, roughly 287 thousand are switches. Thus, even though switching is rare relative to the total number of visits, I still have a sufficiently large number of switches to estimate the true state dependence.

Table 1 presents evidence showing that switching is rare. The table shows the extent of switching conditional on whether the consumer switched a practice or remained loyal to their previous practice in the previous period. The table is calculated for consumers with more than one examination in the data between 2008 and 2017 and each consumer's first observation is removed. A row in the table demonstrates how the states change in the next period, given that the current period's state is the row's state. A column shows what fraction will be at a certain state in the next period out of all consumers who are at the row's state in this period. The first row shows that out of all consumers who were loyal to their practice in the previous visit, 95% remain loyal while 5% switch a practice. The second row shows that 77% of switchers become loyal after the switch, while 23% switch again. The takeaway from the table is that consumers tend to return to their previous practice with a high probability.

3 State Dependence or Preference Heterogeneity?

Both true state dependence and unobserved preference heterogeneity can produce choices with little switching across practices. I disentangle the two competing explanations by controlling for consumers' unobserved preferences towards practices with consumer-practice fixed effects in a multinomial logit using the sufficient statistic approach of Honoré and Kyriazidou (2000). In addition to the fixed effects, I control for two important time-varying covariates affecting practice choice: practices' prices and consumers' distance to the practices. I find evidence of the true state dependence with a magnitude similar to a six-kilometer reduction in distance to the previously visited clinic.

3.1 Controlling for Unobserved Preference Heterogeneity

I model the practice choice using a multinomial logit with consumer-practice fixed effects. The fixed effects approach allows separating the true state dependence from consumers' persistent and unobserved clinic preferences but requires a lot from the data.

For the fixed effect approach of Honoré and Kyriazidou (2000) I assume that consumer i's indirect utility for dental practice j in year t takes the following form

$$u_{ijt} = \alpha p_{jt} + \eta_1 d_{ijt} + \eta_2 d_{ijt}^2 + \gamma 1\{y_{it-1} = j\} + \mu_{ij} + \epsilon_{ijt}$$

where p_{jt} is practice's price, and d_{ijt} is consumer's distance to the practice in time t. The d_{ijt}^2 allows utility to depend non-linearly on distance. The previous period's practice choice is y_{it-1} and $1\{y_{it-1} = j\}$ is an indicator function taking on value one, if the chosen practice in this period is the same as during the previous period. The parameter γ captures the true state dependence and μ_{ij} captures consumer i's time-invariant practice j specific preferences. Finally, ϵ_{ijt} is consumer i's random preference shock that follows an i.i.d. type 1 extreme value distribution. The assumption on the idiosyncratic shock leads to the following multinomial logit expression for the probability that the consumer i chooses the dental practice j in period t at the market m

$$P(y_{it} = j|\mu_{ij}, y_{it-1}, d_{it}, p_{jt}) = \frac{\exp(\alpha p_{jt} + \eta_1 d_{ijt} + \eta_2 d_{ijt}^2 + \gamma 1\{y_{it-1} = j\} + \mu_{ij})}{1 + \sum_{k=1}^{J_m} \exp(\alpha p_{kt} + \eta_1 d_{ikt} + \eta_2 d_{ikt}^2 + \gamma 1\{y_{it-1} = k\} + \mu_{ik})}. \tag{1}$$

Each municipality is a market except the Capital Region where the market is the whole Capital Region, and J_m contains all practices present at the market.

The main challenge in obtaining a consistent estimate of the true state dependence is that unobserved and persistent consumer-practice specific preferences, μ_{ij}, are correlated with the previous period's choice y_{it-1} (Heckman, 1981). Preferences for practices affect which of the practices was chosen in the previous period. If the preferences are left unaccounted for in the model, they enter the error term. This in turn biases the estimates of true state dependence, γ, because the indicator function is correlated with the error term through the previous period's choice.

Another way to see the problem is to note that the true state dependence and consumers' persistent unobserved preference heterogeneity towards practices produce similar choice patterns. When a consumer enters the market, they choose their preferred practice given the prices and distance to all practices. Now suppose they decide to move to another location, that is further from the practice they chose on their first visit. With strong enough preferences towards the practice, the previous practice might still be the most preferred choice, even though there might now be alternatives that are closer. Similarly, with strong enough true state dependence the previous practice might still be the most preferred choice even after moving further from it, for example, because of switching costs, brand loyalty, or habit formation. This observational equivalence also explains why estimates for true state dependence tend to be upward biased when the unobserved preferences are not controlled for.

The Honoré and Kyriazidou (2000) approach solves the main challenge in estimating the true state dependence by allowing the researcher to include the consumer-practice fixed effect into the model. Estimating a nonlinear model with individual specific fixed effects is typically not feasible for data with small and fixed T and large N, like the one that I have, due to the incidental parameter problem (Neyman and Scott, 1948). However, Honoré and Kyriazidou (2000) avoid the incidental parameter problem by coming up with a sufficient statistics S_i such that $P(y_{it} = j|S_i, \mu_{ij}, y_{it-1}, d_{it}, p_{jt}) = P(y_{it} = j|S_i, y_{it-1}, d_{it}, p_{jt})$. Conditional on the sufficient statistics, the probability of choosing an alternative no longer depends on the consumer-practice fixed effects. The downside of the sufficient statistics approach is that it requires a lot from the data.

The sufficient statistic is constructed from two choice sequences of length four, where there must be a practice switch between the second and the third element of the sequence. Consider two choice sequences A and B with four periods indexed by $t = (0, 1, 2, 3)$

$$A = \{y_{i0} = d_{A0}, y_{i1} = w, y_{i2} = l, \;\; y_{i3} = d_{A3}\}$$
$$B = \{y_{i0} = d_{B0}, y_{i1} = l, \;\; y_{i2} = w, y_{i3} = d_{B3}\}.$$

Sequences must be such that if a consumer with the sequence A chooses w in period one and l in period two, then the sequence B must be such that l is chosen in period one and w is chosen in period two. Choices in periods zero and three are not restricted, $d_{A0}, d_{A3}, d_{B0}, d_{A0} \in J_m$. In my analysis, the sequence can be for example from 2008 to 2012, or from 2010 to 2015 such that consumers did not visit during 2011. However, I do not allow gaps between the two middle periods.

Honoré and Kyriazidou (2000) show that the fixed effects disappear when conditioning on observing either the sequence A or the sequence B and on any covariates included in the model being constant between periods two and three. In my case, the requirement on covariates means that I can only use observations where prices and distance to the practice do not change between periods two and three. For distance, this merely decreases the sample size as I drop those consumers who move between periods two and three. However, the case with prices is more complicated. Each multinomial logit expression for the probability of choosing an alternative contains all prices for all the firms in the market. Moreover, it is very unlikely that all firms in a market would keep their prices constant for two years.

The solution that Honoré and Kyriazidou (2000) have for including prices, or more generally a continuous variable, in the model is to kernel weight the likelihood so that more weight is given to observations where the difference in prices is smallest between the last two periods two and three. To obtain the weights, I use a normal kernel with a bandwidth $h = c\sigma_n n^{-1/(4+J_m)}$, which is equal to Scott's rule of thumb for $c = 1.06$. n is the number of sequences I use to estimate the fixed effects model. σ_n is the standard deviation of price differences between sequence-specific periods 2 and 3 for all practices in the markets J_m and time periods that I have in my fixed effects model estimation sample. In my sample, σ_n is 4.76.

The choice of the bandwidth is more important than the choice of the kernel. The optimal bandwidth for the standard normal kernel with k dimensions is $h = cn^{-1/(4+k)}$ for some c. My inclusion of σ_n into the formula can be interpreted in two ways. The first interpretation is that I allow c to depend on σ_n. The second interpretation is that instead of using a standard normal kernel I am using a kernel with unknown standard deviation and hence the standard deviation of the price differences enters the formula.

The downside of including prices in the Honoré and Kyriazidou (2000) model is that the converge rate of the estimator is lower than the standard $\sqrt{n}$ rate when a continuous variable is included. Assuming a standard normal kernel and the optimal bandwidth, the rate of convergence is $n^{\frac{-1}{4+k}}$ where the k is the dimension of the kernel or equivalently the number of continuous covariates in the model (Aeberhardt and Davezies, 2012).

These undesirable properties of the model when price is included as a covariate motivate my addition of the σ_n into the formula for the kernel's bandwidth. Honoré and Kyriazidou (2000) show in their simulations that with a small sample and with a small bandwidth the estimator produces estimates that are on average upwards biased, while with too large bandwidth the estimates are on average biased downwards. This is because while the estimator is consistent, it is biased. Using a bandwidth without the standard deviation and with $c = 1.06$ leads to dropping a little less than half of the choice sequences in the sample that can be used when prices are not included. This happens because the kernel gives some sequences zero weight when the standard deviation is omitted. When the standard deviation is included, all choice sequences get positive weight. Hence, including the standard deviation in the bandwidth should help balance the trade-off between the small sample bias and the bias from too large bandwidth. Moreover, having the same sample in the model with and without prices facilitates comparison across the models.

Conditioning on observing either an A sequence or a B sequence and omitting sequences where the consumer moves between the second and the third period, we can estimate the parameters of the multinomial logit specification in equation 1 using maximum likelihood. The likelihood function I take to the data is

$$
\begin{aligned}
&P(A|d_i, p_i, A \cup B, d_{i2} = d_{i3}) \\
&= \sum_{i=1}^{N} \sum_{w \neq l} 1\{y_{iw1} + y_{il2} = 1\} \times \{d_{i2} - d_{i3} = 0\} \times K\left(\frac{p_{2m} - p_{3m}}{h}\right) \\
&\times \ln\left(\frac{\exp[\alpha(p_{iw1} - p_{iw2} + p_{il2} - p_{il1}) + \eta_1(d_{iw1} - d_{iw2} + d_{il2} - d_{il1}) + \eta_2 d_i^2 + \gamma b_i]^{y_{iw1}}}{1 + \exp[\alpha(p_{iw1} - p_{iw2} + p_{il2} - p_{il1}) + \eta_1(d_{iw1} - d_{iw2} + d_{il2} - d_{il1}) + \eta_2 d_i^2 + \gamma b_i]}\right)
\end{aligned}
\tag{2}
$$

where $b_i = 1\{y_{i0} = w\} + 1\{y_{i3} = l\} - \Big(1\{y_{i0} = l\} + 1\{y_{i3} = w\}\Big)$ and $d_i^2 = (d_{iw1}^2 - d_{iw2}^2 + d_{il2}^2 - d_{il1}^2)$. The indicator variable y_{iw1} takes on value one, if the consumer with the sequence i visited practice w in period one, which means that the sequence i is of A-type. The variable containing variation on consumers' switching behavior, b_i, takes on a large value when the consumer chooses the same practice in period zero as in period one and when the consumer chooses the same practice in period three as in period two. So the variation on b_i comes from switches between periods zero and one and between periods two and three. p_{2m} is a column vector with prices for all practices at the market J_m in period two and for the J_m dimensional kernel, K, I assume that the dimensions are independent. The likelihood function reveals that the panel dimension of the data has disappeared. This is because Honoré and Kyriazidou (2000) approach collapses the sequences of choices into a cross-section to get rid of the consumer-practice fixed effects. Finally, note that the Honoré and Kyriazidou (2000) approach does not require specifying an outside good.

The identifying variation for the parameters on distance and on price comes from how the number of A sequences changes relative to the B sequences when the covariates for the alternatives chosen in the two middle periods, w and l, change between periods one and two. Remember that for a sequence of type A, practice w was chosen in period one and practice l in period two, and for type B, l was chosen in period one, and w was chosen in period two. What the log-likelihood function in the equation 2 does is that it compares the changes in distance and prices for practices w and l between periods one and two to the number of sequences A and to the number of sequences B. So the identifying variation for the distance parameter and for the price parameter is analogous to a binary logit where the dependent variable is whether the choice sequence is of A or of B type. In this binary logit, if the prices change so that the price of the practice w increases between periods one and two while the practice l's price decreases between the periods, then we should observe more switches from w to l than switches from l to w if consumers dislike higher prices.

The variation for the distance and for the price consists of a sum of two differences. Let us consider how the expression for the distance, $d_{iw1} - d_{iw2} + d_{il2} - d_{il1}$, works. First, the distance to the practice w in period two is subtracted from the distance to the practice w in period one. Then the distance to practice l in period one is subtracted from the distance to practice l in period one. Finally, these differences are added together. Consider an example where a consumer moves between periods one and two. Suppose that the consumer moves two kilometers further from the practice w and three kilometers closer to the practice l. Then $d_{iw1} - d_{iw2} = -2$ and $d_{il2} - d_{il1} = -3$, and thus $d_{iw1} - d_{iw2} + d_{il2} - d_{il1} = -5$. In this case, we would expect that the consumer chose the practice w in period one and switched to the practice l in period two if consumers dislike longer distances to their practice. If we observe that many consumers who moved in exactly this way are of sequence A, so chose w and then switched to l and not the other way around, then the maximization procedure should give a negative sign on the distance covariate.

3.2 Honoré and Kyriazidou Estimation Sample

The required restrictions shrink the panel data from five and half million observations down to a cross-section of 4635 observations, each containing a sequence of three or four visits. Fortunately, the consumers in the restricted sample are similar in observable characteristics to consumers in the full sample.

To implement the Honoré and Kyriazidou (2000) approach, I need to find sets of consumers such that in each set I have at least one consumer with a choice sequence of type A and at least one consumer with a choice sequence of type B. Choice sequences of type A require that a consumer has chosen an alternative w in period one and alternative l in period two. The corresponding choice sequence B requires that a consumer has chosen alternatives w and l in the opposite order, so l in period one and w in period two. The alternatives can be any two practices within a market. I end up with many sets of A and B sequences such that each set contains at least one incidence of both types of sequences. A set of A and B sequences is defined by which practices the w and l are and in which years the sequences' visits occur. In a set of A and B sequences, all period one visits must have occurred in the same year, and the same holds for all periods, so also for periods zero, two, and three.

The second restriction is that I need in total four observations for each sequence, but for consumers who enter the market during my sample, I only need three observations. Raval and Rosenbaum (2018) shows that it is possible to use the Honoré and Kyriazidou (2000) approach with three observations when the researcher observes a consumer beginning their choice process during the sample. In this case, the period zero observation is not needed because the researcher knows that the period one observation was the first choice event for the consumer. If the period one observation is not the first choice event for the consumer, and the researcher still attempts to use the Honoré and Kyriazidou (2000) approach with only observations one, two, and three, the researcher would face the initial condition problem of assuming what alternative was chosen in period zero visit.

Using choice sequences of length three increases the sample size, but it also has some downsides. First, using only three observations the variable containing the information about switching, b_i, has less variation. The variable loses variation that comes from whether the consumer chose the same alternative in period zero as in period one. Second, Raval and Rosenbaum (2018) only used their approach with discrete variables. Hence, I am the first researcher to use the three observations with the kernel. Nonetheless, the results are similar whether I use sequences of four or three observations.

Table 2 presents summary statistics on the Honoré and Kyriazidou estimation sample. The final sample used in the Honoré and Kyriazidou estimation is significantly smaller than the full sample of five and half million observations, containing only 4635 choice sequences, amounting to 17 517 consumer-year observations. The two last rows of the table show that roughly three-fourths of the sequences are four visits long, while the rest are three visits long. The first three rows of Table 2 show the variation used to estimate the effect of the state dependence, distance, and prices on the health care practice choice. The variable for state dependence, b_i, varies from minus two to two depending on whether the consumer switched their practice between periods zero and one, and between periods two and three. If there were no switches, then the variable takes on value two, while if there were two switches, then the variable takes on value minus two. The mean of the state dependence variable is 0.970, which implies that in most of the two-period pairs, there is no switching. Price, Δp_j, has a good amount of variation with a standard deviation of 4.86. Unfortunately, most consumers do not move between periods 1 and 2 so the number of sequences with variation in distance covariate is low, which explains why the 25th percentile for Δd_{ij} is 0.

The small sample size raises a concern whether the Honoré and Kyriazidou sample is representative of the full sample. Table 3 addresses the concern by showing the average patient, average dental practice, and average market characteristics of the full sample next to the average patient, dental practice, and market characteristics of the Honoré and Kyriazidou samples. The first column shows the average characteristics of the full sample. The second column, HK Sample, shows the average characteristics of the Honoré and Kyriazidou sample. The third column, KW HK Sample, presents averages for the Honoré and Kyriazidou samples but weighted with the kernel weights.

Table 3 Panel A shows that the consumers in HK and KW HK sample The table shows that the samples are similar in terms of average consumer characteristics. The main differences are that KW HK Sample has lower earnings and that consumers in HK sample have on average lower distance to dental practice than the full sample's consumers. Consumers in HK and KW HK samples are also more likely to visit a dentist, which is not surprising considering that consumers with many visits are more likely to qualify for the estimation sample.

Table 3 Panel B shows that my estimation sample contains one-fourth of all dental practices. Practices in the full and HK samples are similar on average, but practices in HK sample are on average larger as on average they perform more dental examinations and employ more dentists than practices in the full sample. Practices in the KW HK sample are substantially smaller on average than the practices in either the full or KW sample. This likely reflects the fact that observations in markets with large price changes across years and many different practices get a lower weight than practices in markets with smaller price changes and a low number of practices. The former markets are likely to be more urban and the latter are likely to be more rural. Thus in KW HK sample rural markets have a higher weight.

Table 3 Panel C shows average market characteristics in different samples. The HK and KW HK samples contain 70 of all 173 markets. The markets in the HK sample appear to be more urban as they have on average the lowest Herfindahl-Hirschman index (HHI), and the highest number of practices, dentists, and patients. Markets in the KW HK sample are more rural on average, with the highest HHI and the lowest number of practices, dentists, and patients on average. The main takeaway from this section is that consumers in HK and in KW HK sample are similar to the full population, while the practices in HK sample are larger and practices in KW HK sample are smaller than the practices in the full sample. Markets in HK sample are on average more urban while markets in KW HK sample are less urban than markets in the full sample.

3.3 Results and Discussion

Table 4 presents estimation results for three model specifications: the first one has only the state dependence as a covariate, the second one adds distance, and the third adds both distance and prices as well as the kernel. The Model 1 produces a state dependence estimate of

0.171 and the estimate is statistically significant at the conventional levels. Adding distance to the model leaves the state dependence estimate practically unchanged. The estimate for the distance is statistically significant -0.143 with the expected sign, meaning that consumers prefer practices that are closer to them. The estimate for the second power of distance is positive, meaning that the disutility of distance increases at a decreasing rate as distance increases. However, the estimate is not statistically significant. Moving on to Model 3 the estimate for the state dependence increases to 0.220, and the estimate is no longer statistically significant even at a 10% level. The estimate for the distance becomes smaller in absolute value and is now -0.031 and ceases to be statistically significant. The second power of distance changes sign to negative but is not statistically significant. The estimate for price is negative as expected, but not statistically significant.

The state dependence estimate can change across Model 2 and Model 3 because of two differences: omitted variable bias in Model 2 due to not controlling for price, or because of the different weighting of observations due to kernel in Model 3. Including prices in the model is important, because changes in prices over time affect the probability of choosing a practice, and they are correlated with the previous period's choice through a similar mechanism as today's choice. Omitting price from the model might hence expose the model to omitted variable bias. On the other hand, the kernel in Model 3 gives different weights to the observation than the Model 2. It could be that the weights in Model 3 decrease the power so that the standard errors of Model 3 increase relative to Model 2 and thus I lose the power to identify even quite large effects.

Table 5 presents robustness checks for different values of bandwidths. Smaller values of c result in more stringent weights in the sense that markets with larger price increases and higher number of practices get less weight. Larger c results in more equal weights. The state dependence parameter is quite robust to different values of c, and the coefficient only increases with smaller values of c. Larger values of c lead to the same coefficient as in Model 2 without the price. Moreover, of the four estimates, two are statistically significant at 10% level. Finally, note that reducing c reduces the sample size, as with small values of c the kernel gives zero weight to some observations. I conclude from these estimates that the loss of significance between the state dependence estimate from Models 2 and 3 is due to a loss of power when using weights.

I consider the Model 2 estimate of the state dependence to be my preferred estimate for the true state dependence. To get a sense of how important the true state dependence is, I divide the state dependence estimate with the estimate for marginal utility of distance: $\eta_1 + 2\eta_2\overline{d_{jt}}$, where $\overline{d_{jt}}$ is the average distance in the whole population, 16.15 km. This produces a value of approximately three and a half, meaning that visiting a practice in the previous period increases the probability of choosing the practice in this period as much as reducing the distance to the practice by three and a half kilometers. This is a sizable effect, as the mean distance to a chosen practice in the full sample is 16 kilometers.

3.4 Limitations

My empirical approach of estimating the extent of true state dependence has some limitations. The first is that I am not able to control for any time-fixed effects due to using Honoré and Kyriazidou (2000) approach. The second is that there is still some uncertainty about how well the results from the estimation sample can be generalized to the full population using private dental care. Thirdly, I am not able to address the fact that price is endogenous. Even though I am able to control for practices' time-invariant characteristics, and consumer-practice time-invariant characteristics, any time-varying sources of endogeneity remain a threat to the identification.

I assume that consumer demand is static. However, there are papers where switching costs imply that consumers' demand is dynamic, like Nosal (2012) and Shcherbakov (2016). The reason why switching costs make the demand dynamic in these articles is that the papers study markets where the product characteristics are changing rapidly: Medicare Advantage plans and the paid television industry in the US. Characteristics of the dental practices are quite stationary year to year, and hence I argue that a static demand model is sufficient for studying choice frictions in this market. Moreover, my focus is on controlling for preference heterogeneity in a flexible way. My ability to do this would be severely limited if I adopted a dynamic demand framework.

4 Conclusion.

I find evidence of choice frictions in dental care practice choice. I rule out consumers' unobserved persistent preference heterogeneity towards dental practices as an explanation for the low number of practice switches in the data by controlling for consumer-practice-specific fixed effects. The magnitude of the true state dependence arising from the choice frictions is similar to a three-and-a-half kilometer decrease in distance to the previous period's practice. This is a sizable effect considering that the average distance from a consumer to the practice they are visiting is 16 kilometers in the full sample.

My results indicate that choice frictions are important in the primary health care industry, and should be taken into account as one of the reasons why health care markets might fail. I am not able to differentiate between different sources of choice friction. However, whether it might make sense to try to reduce choice frictions with policy measures depends on the choice frictions in question future work should find out what are the main mechanisms that give rise to the choice frictions in the primary health care context. This information would be especially valuable to policymakers.

References

Abaluck, J. and J. Gruber (2011). Choice Inconsistencies Among the Elderly: Evidence from Plan Choice in the Medicare Part D Program. *American Economic Review 101*(4), 1180–1210.

Aeberhardt, R. and L. Davezies (2012). Practical Guidelines for the Estimation and Inference of a Dynamic Logistic Model with Fixed-Effects. *Economics Letters 115*(2), 300–304.

Anell, A., J. Dietrichson, L. M. Ellegård, and G. Kjellsson (2021). Information, Switching costs, and Consumer Choice: Evidence from Two Randomised Field Experiments in Swedish Primary Health Care. *Journal of Public Economics 196*, 104–390.

Arrow, K. (1963). Uncertainty and te Welfare Economcis of Medical Care. *The American Economic Review 53*(5), 941–973.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica 63*(4), 841–890.

Buchmueller, T., S. Miller, and M. Vujicic (2016). How Do Providers Respond to Changes in Public Health Insurance Coverage? Evidence from Adult Medicaid Dental Benefits. *American Economic Journal: Economic Policy 8*(4), 70–102.

Chintagunta, P., E. Kyriazidou, and J. Perktold (2001). Panel data analysis of household brand choices. *Journal of Econometrics 103*(1-2), 111–153.

Choi, M. K. (2011). The Impact of Medicaid Insurance Coverage on Dental Service Use. *Journal of Health Economics 30*(5), 1020–1031.

Dahl, G. B. and S. J. Forbes (2023). Doctor switching costs. *Journal of Public Economics 221*, 104858.

Decker, S. L. and B. J. Lipton (2015). Do Medicaid Benefit Expansions Have Teeth? The Effect of Medicaid Adult Dental Coverage on the Use of Dental Services and Oral Health. *Journal of health economics 44*, 212–225.

Dubé, J.-P., G. Hitsch, and P. Rossi (2009). Do Switching Costs Make Markets Less Competitive. *Journal of Marketing Research 46*(3), 435–445.

Dubé, J.-P., G. J. Hitsch, and P. E. Rossi (2010). State Dependence and Alternative Explanations for Consumer Inertia. *The RAND Journal of Economics 41*(3), 417–445.

Erdem, T. and B. Sun (2001). Testing for Choice Dynamics in Panel Data. *Journal of Business and Economics Statistics 19*(2), 142–152.

Farrell, J. and P. Klemperer (2007). Chapter 31 Coordination and lock-in: Competition with switching costs and network effects. In *Handbook of Industrial Organization*, Volume 3, pp. 1967–2072. Elsevier.

Feng, J. (2022). History-Dependence in Drug Demand: Identification and Implications for Entry Incentives. *The Review of Economics and Statistics*, 1–45.

Giannobile, W., T. Braun, A. Caplis, L. Doucette-Stamm, G. Duff, and K. Kornman (2013). Patient stratification for preventive care in dentistry. *Journal of Dental Research 92*(8), 694–701.

Granlund, D. (2021). A New Approach to Estimating State Dependence in Consumers' Brand Choices Applied to 762 Pharmaceutical Markets. *The Journal of Industrial Economics 69*(2), 443–483.

Grzybowski, L. and A. Nicolle (2021). Estimating Consumer Inertia in Repeated Choices of Smartphones. *The Journal of Industrial Economics 69*(1), 33–82.

Handel, B. R. (2013). Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts. *American Economic Review 103*(7), 2643–2682.

Heckman, J. J. (1981). Heterogeneity and State Dependence. In S. Rosen (Ed.), *Studies in Labor Markets*, Chapter 3, pp. 91–140. University of Chicago Press.

Heiss, F., D. McFadden, J. Winter, A. Wupperman, and B. Zhou (2021). Inattention and Switching Costs as Sources of Inertia in Meidcare Part D. *American Economic Review 68*(9), 2737–2781.

Honoré, B. and E. Kyriazidou (2000). Panel Data Discrete Choice Models with Lagged Dependent Variables. *Econometrica 68*(4), 839–874.

Hortaçsu, A., S. A. Madanizadeh, and S. L. Puller (2017). Power to choose? An analysis of consumer inertia in the residential electricity market. *American Economic Journal: Economic Policy 9*(4), 192–226.

Irace, M. (2018). Patient Loyalty in Hospital Choice: Evidence from New York. Becker Friedman Institute for Economics Working Paper 2018-52, University of Chicago.

Janssen, A. (2022). Generic and Branded Pharmaceutical Pricing: Competition under Switching Costs.

Kong, X., J.-P. Dubé, and Ø. Daljord (2022). Non-parametric Estimation of Habitual Brand Loyalty. Working paper.

Kwok, J. (2024). How do primary care physicians influence healthcare? evidence on practice styles and switching costs from medicare. Working paper.

Luco, F. (2019, may). Switching costs and competition in retirement investment. *American Economic Journal: Microeconomics 11*(2), 26–54.

MacKay, A. and M. Remer (2022). Consumer Inertia and Market Power. Working Paper 19-111, Harvard Business School.

Mehiläinen (2022). Dental check-up. https://www.mehilainen.fi/en/hammasmehilainen/dentalcheckup. Online; accessed 21 November 2022.

Neyman, J. and E. Scott (1948). Consistent Estimates Based on Partially Consistent Observations. *Econometrica*, 1–32.

Nosal, K. (2012). Estimating Switching Costs for Medicare Advantage Plans. *Unpublished manuscript, University of Mannheim.*

Nurminen, M., J. Blomgren, and H. Mikkola (2021). Socioeconomic Differences in Utilization of Public and Private Dental Care in Finland: Register-based Evidence on a Population Aged 25 and Over. *PloS one 16*(8).

Pakes, A., J. Porter, M. Shepard, and S. Calder-Wang (2021). Unobserved heterogeneity, state dependency, and health plan choices. Working Paper 29025, National Bureau of Economic Research.

Polyakova, M. (2016). Regulation of insurance with adverse selection and switching costs: Evidence from medicare part D. *American Economic Journal: Applied Economics 8*(3), 165–195.

Raval, D. and T. Rosenbaum (2018). Why Do Previous Choices Matter for Hospital Demand? Decomposing Switching Costs from Unobserved Preferences. *The Review of Economics and Statistics 100*(5), 906–915.

Sabety, A. (2022). The Value of Relationships in Health Care. Technical report.

Shcherbakov, O. (2016). Measuring consumer switching costs in the television industry. *Journal of Economics 47*(2), 366–393.

The Finnish Institute for Health and Welfare (2021a). Health Expenditure and Financing 2019. Statistical Report 15/2021, The Finnish Institute for Health and Welfare.

The Finnish Institute for Health and Welfare (2021b). Sotkanet Indicator Bank Indicator 2498. https://sotkanet.fi/sotkanet/en/taulukko/?indicator=s3YxsjZzBAA=&region=s07MBAA=&year=sy5ztrbU0zUEAA==&gender=t&abs=f&color=f&buildVersion=3.1.1&buildTimestamp=202211091024. Online; accessed 18 November 2022.

The Finnish Institute for Health and Welfare (2021c). Sotkanet Indicator Bank Indicator 2918. https://sotkanet.fi/sotkanet/en/taulukko/?indicator=s3YxsjZzBAA=&region=s07MBAA=&year=sy5ztrbU0zUEAA==&gender=t&abs=f&color=f&buildVersion=3.1.1&buildTimestamp=202211091024. Online; accessed 18 November 2022.

Tilipman, N. (2022). Employer Incentives and Distortions in Health Insurance Design: Implications for Welfare and Costs. *American Economic Review 112*(3), 998–1037.

Wing, C. and A. Marier (2014, mar). Effects of occupational regulations on the cost of dental services: Evidence from dental insurance claims. *Journal of Health Economics 34*(1), 131–143.

Yeo, J. and D. P. Miller (2018). Estimating switching costs with market share data: An application to medicare part d. *International Journal of Industrial Organization 61*, 459–501.

5 Tables

Table 1: Markov Transition Matrix for Visiting the Same practice

	Loyal	Switched
Loyal	0.95	0.05
Switched	0.77	0.23

Note: A row should be interpreted as from where and a column as to where. The table is calculated for consumers with more than one examination in the data between 2008 and 2017. One observation is one treatment episode. Each treatment episode begins with an examination and contains several procedures performed over several visits. Examinations are used to proxy for one treatment episode.

Table 2: Summary Statistics for Honoré and Kyriazidou Estimation Sample

Statistic	N	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)
b_i	4,635	0.970	1.038	0	1	2
Δd_{ij}	4,620	−0.009	0.867	0.000	0.000	0.000
Δp_j	4,635	−1.243	4.863	−2.446	−0.559	0.638
Share sequence length 4	4,635	0.779	0.415	1	1	1
Share sequence length 3	4,635	0.221	0.415	0	0	0

Note: $b_i = 1\{y_{i0} = w\} + 1\{y_{i3} = l\} - \big(1\{y_{i0} = l\} + 1\{y_{i3} = w\}\big)$.

Table 3: Comparison Between the Full Sample and Honoré and Kyriazidou Sample

Panel A: Consumer Characteristics

	Full sample	HK sample	KW HK Sample
Mean Age	53.81	54.00	55.21
Share Women	0.57	0.62	0.61
Mean Earnings (euro 1 000s)	35.03	37.11	27.74
Mean Capital income (euro 1 000s)	3.57	3.92	3.19
Mean Distance to provider (km)	16.15	10.26	14.36
Mean Number of visits in the full sample	5.30	7.05	6.66
N consumer-year observations	5, 550, 642	17, 517	17, 517

Panel B: Dental Practice Characteristics

	Full sample	HK sample	KW HK Sample
Mean Real price	60.73	60.65	54.28
Mean Within year Std. Dev. of real price	7.59	7.28	4.07
Mean Dentist age	49.46	47.63	49.86
Mean Dentists' graduation year	1, 989.49	1, 991.73	1, 986.82
Share Women dentists	0.57	0.59	0.40
Share Specialist dentists	0.08	0.08	0.02
Share Extra time reimbursement	0.0004	0.001	0.0001
Share Extra specialist reimbursement	0.04	0.03	0.01
Mean N Examination patients per year	1, 279.00	1, 808.87	598.31
Mean Number of dentists per month	4.25	6.32	1.78
N practices	2, 316	630	630

Panel C: Market Characteristics

	Full sample	HK sample	KW HK Sample
Mean Real price	60.73	61.11	54.05
Mean Within year Std. Dev. of real price	19.99	17.53	6.73
Examination HHI	1, 735.17	1, 032.55	4, 219.01
Mean Number of Practices	60.11	76.12	4.01
Mean Number of Dentists	151.67	189.47	5.98
Mean Number of Patients	28, 515.67	37, 306.89	1, 751.20
Mean Patients per dentist	258.73	244.37	338.43
N Municipalities	173	70	70

Note: The Panel A compares the average consumer-year observation in the full sample to the average consumer-year observation in the Honoré and Kyriazidou sample. KW HK Sample is HK sample weighted with the kernel such that the weights sum up to one and multiplied with the total number of observations in HK sample.

Table 4: Results from the Honoré and Kyriazidou (2000) Fixed Effects Model

	Model 1	Model 2	Model 3
γ Loyalty	0.171^{***}	0.168^{***}	0.220
	(0.021)	(0.021)	(0.177)
η_1 Distance (km)		-0.143^{**}	-0.031
		(0.064)	(0.331)
η_2 Distance^2		0.003	-0.005
		(0.002)	(0.029)
α Price			-0.061
			(0.100)
Kernel	No	No	Yes
N Sequences	4635	4620	4620

Note: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are in the parentheses. In Model 3 c is set to 1.06, so the bandwidth is Scott's rule of thumb, $c\sigma_n n^{1/(4+J_m)}$. Standard errors for Model 1 and Model 2 are obtained from the Fisher information matrix. For Model 3 the standard errors are obtained using the approach by Aeberhardt and Davezies (2012) due to the estimator's non-standard rate of convergence. The different sample sizes between Model 1 and Model 2 and 3 are due to missing observations for distance.

Table 5: Robustness of the Model with Kernel to Different Bandwidths with Different c

$c =$	0.5×1.06	1.06	2×1.06	4×1.06
γ Loyalty	0.502	0.220	0.216^{*}	0.162^{*}
	(0.310)	(0.177)	(0.131)	(0.093)
η_1 Distance (km)	0.955	-0.031	-0.136	-0.166
	(1.198)	(0.331)	(0.300)	(0.295)
η_2 Distance^2	0.080	-0.005	0.002	0.003
	(0.061)	(0.029)	(0.026)	(0.026)
α Price	-0.183	-0.061	-0.052	-0.049
	(0.151)	(0.100)	(0.070)	(0.052)
Kernel	Yes	Yes	Yes	Yes
N Sequences	4164	4620	4620	4620

Note: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. The bandwidth is $c\sigma_n n^{1/(4+J_m)}$, with $\sigma_n = 4.76$. When $c = 1.06$, the bandwidth is Scott's rule of thumb.