department of economics
"S. Cognetti de Martiis"

Via Po, 53 – 10124 Torino (Italy)
Tel. (+39) 011 6704043 - Fax (+39) 011 6703895
URL: http//www.de.unito.it

WORKING PAPER SERIES

Longitudinal analysis of generic substitution

Dalen D. M., Locatelli M. and Strøm S.

Dipartimento di Economia "S. Cognetti de Martiis"

Working paper No. 13/2010

Università di Torino

12/06/2010

Longitudinal analysis of generic substitution.[1]

by

Dag Morten Dalen, Marilena Locatelli, Steinar Strøm

Abstract

Using an extensive longitudinal dataset extracted from the Norwegian Prescription Database (NorPD) containing all prescriptions written in the period January 2004 to June 2007, we selected two particular drugs (chemical substances) used against cholesterol. The two brand-name products on the Norwegian markets were Provachol (atc code C10AA03) and Zocor (atc code C10AA01). The generics are Provastatine and Simastatine. The model accounts for taste persistence and is estimated on panel data. We find that prices have a negative impact on transitions in the sense that an increase in the brand price will reduce the transition from generics to brand and likewise an increase in the generic price will reduce the transition from brand to generics.

JEL: C35, I 18, L 65

Keywords: Generics, substitution, microdata, random utility model, longitudinal data

Addresses:

Dag Morten Dalen, BI Norwegian School of Management, email: dag.m.dalen@bi.no

Marilena Locatelli, University of Turin, Department of Economics, Turin, email: marilena.locatelli@unito.it

Steinar Strøm, The Frisch Centre, Oslo, email: steinar.strom@econ.uio.no

[1] We wish to acknowledge the services of the NorPD (and particularly of Kari Furu) in providing data. This research is supported by grants from the Norwegian Research Council.
Financial support to M. Locatelli from The Frisch Centre, Oslo is gratefully acknowledged.

1. Introduction

In Dalen et al (2010) we estimated the choice between brand-name and generic drugs based on cross-section data. We extracted the entire population of prescriptions in February 2004 and 2006 on 23 different chemical substances. In February 2004 we had 102 201 observations and in February 2006 we had 210 877 observations. The observations gave us the choice of brand or generics among these patients in these two cross-sections. From the estimated model we derived price elasticities which were the elasticities of the brand products with respect to the brand price. The average of these elasticities was -0.36 in 2004.

In the present paper we exploit the longitudinal dimension of the data and estimate a dynamic model on monthly observations from May 2004 until June 2007 of drug choices for 109 patients in Norway. From the model we derive transition probabilities that give the transition from brand-name drug to generics and vice versa. We selected only one drug; a drug used against cholesterol. The two brand-name products on the Norwegian markets for *statines* in the period May 2004 to June 2007 were *Provacho*l (atc code C10AA03) and *Zocor* (atc code C10AA01). The generics are *Provastatine* and *Simastatine*. From the model we derived elasticities of the probabilities of shifting from brand to generics with respect to the price of generics and of the probabilities of shifting from generics to brand with respect to the brand price. The average of the elasticities over patients and periods were -0.27 and -0.46 respectively which are not that different from the estimates of the price elasticity derived from the cross-section estimates referred to above which also covered not only statines but 22 other substances.

In addition to the expected price effects we found that the older a male doctor is the more likely it is that he continues to prescribe the brand-name product. The dynamic model allows for taste persistence and the correlation of is calculated across patients and across time.

The paper is organized as follows. Section 2 presents the model. Section 3 gives the data, estimates are given in Section 4 and Section 5 concludes.

2. The model

The model we employ is based on a dynamic choice model developed by Dagsvik (2002). Let $U_{nj}(t)$ denote the utility of patient n of using drug j at time t. j= B(brand-name), G(generics). Let B_{nt} be the choice set. We will assume that $\{U_{nj}(t), j \in B_{nt}\}$ is a random utility process. Let $\{v_{nj}(t) + \varepsilon_{nj}(t)\}$ be the period-specific utility in contrast to $U_{nj}(t)$ which are utilities that account for "taste-persistence". The $\varepsilon_{nj}(t)$ are assumed to be independent of $v_{nj}(t)$ and they are assumed to be iid extreme value distributed, that is $\Pr(\varepsilon_{nj}(t) \le x) = \exp(-\exp(x))$.

The model extends the common logit model to deal with correlation in preferences or rather *taste persistence*. It should be noted that this is not the same as *state dependence*. With the latter the choice you have made in the past has a direct impact on the current choices. This is not the case here; the assumption is simply that preferences may be correlated. In Dagsvik (2002) it is shown that

(1) $$U_{nj}(t) = \max(U_{nj}(t-1) - \theta, v_{nj}(t) + \varepsilon_{nj}(t))$$

The coefficient θ may be interpreted as a preference discount factor:

If **θ=0** there is a complete strong taste persistence, and if **θ=∞** there is no taste persistence at all and $U_{nj}(t) = v_{nj}(t) + \varepsilon_{nj}(t)$.

The expected value of $U_{nj}(t)$ is given by

(2) $$EU_{nj}(t) = \ln \sum_{r=t_0}^{t} \exp(v_{nj}(r) - (t-r)\theta)$$

or

(3) $$\exp(EU_{nj}(t)) = \sum_{r=t_0}^{t} \exp(v_{nj}(r) - (t-r)\theta)$$

To calculate correlation across utilities it is convenient to calculate correlation of a monotone transformation of the utilities:

(4) $$\mathrm{corr}(\exp(-U_{nj}(s)), \exp(-U_{nj}(t))) = \frac{\exp(EU_j(s))}{\exp(EU_j(t))} e^{-(t-s)\theta} \text{; for } s \leq t$$

We observe that if covariates are constant over time the correlation from t to t-1 is approximately equal to $e^{-\theta}$.

As shown in Dagsvik (2002) the model can be employed to yield transition probabilities, which in our case will be between brand-name products and generics. Thus the transition probabilities are the following:

Q_{nBGt} = probability that patient *n* transit from **B**rand-name drug in period t-1 to **G**enerics in period t

Q_{nBBt} = probability that patient n stay on **B**rand-name drug in period t-1 and in period t

Q_{nBBt}= 1-Q_{nBGt}

Q_{nGBt}= probability that patient n transit from **G**enerics in period t-1 to **B**rand-name drug in period t

Q_{nGGt}= probability that patient n stay on **G**eneric in period t-1 and in period t

Q_{nGGt}= 1-Q_{nGBt}.

The transition probabilities have the following structure:

(5) $$Q_{nBGt} = \frac{\exp(v_{nGt})}{\sum_{r=t_0}^{t} \left[\exp(-(t-r)\theta_n)\right]\left[\exp(v_{nGr}) + \exp(v_{nBr})\right]}$$

(6)

$$Q_{nBBt} = 1 - Q_{nBGt} = \frac{\exp(v_{nBt}) + \sum_{r=t_0}^{t-1}\left[\exp(-(t-r)\theta_n)\right]\left[\exp(v_{mGr}) + \exp(v_{nBr})\right]}{\sum_{r=t_0}^{t}\left[\exp(-(t-r)\theta_n)\right]\left[\exp(v_{nGr}) + \exp(v_{nBr})\right]}$$

(7) $$Q_{nGBt} = \frac{\exp(v_{nBt})}{\sum_{r=t_0}^{t}\left[\exp(-(t-r)\theta_n)\right]\left[\exp(v_{nGr}) + \exp(v_{nBr})\right]}$$

(8)

$$Q_{nGGt} = 1 - Q_{nGBt} = \frac{\exp(v_{nGt}) + \sum_{r=t_0}^{t-1}\left[\exp(-(t-r)\theta_n)\right]\left[\exp(v_{nGr}) + \exp(v_{nBr})\right]}{\sum_{r=t_0}^{t}\left[\exp(-(t-r)\theta_n)\right]\left[\exp(v_{nGr}) + \exp(v_{nBr})\right]}$$

The deterministic part of the utility function, v_{jnt}, j=B,G is assumed to depend linearly on the price of the drug, age and gender of patient. Because of the loyalty among patients and doctors we expect that θ_n will have a low value indicating strong taste persistence. θ_n may depend on characteristics such as age and gender of doctors and patients. However, here we assume it to be a constant.

t_0=date of entry of the drug to the market. Because the data we use are detailed register data that started in January 2004, t_0 is set equal to this date.

The model is estimated by a standard maximum likelihood procedure. The likelihood is:

(9) $$L = \prod_n \prod_t Q_{BGnt}^{y_{nt}} (1 - Q_{BGnt})^{1-y_{nt}} Q_{GBnt}^{z_{nt}} (1 - Q_{GBnt})^{1-z_{nt}}$$

where:

(10) $$y_{nt} = \begin{cases} 1 \text{ if transition from Brand to Generic} \\ 0 \text{ otherwise} \end{cases}$$

(11) $$z_{nt} = \begin{cases} 1 \text{ if transition from Generic to Brand} \\ 0 \text{ otherwise} \end{cases}$$

We assume that the deterministic part of the utility function depends on the price of the drug, and the interaction between age and gender of both patient and doctor. We expect that price has a negative impact on demand. Furthermore we expect that male patient, in particular when they are getting older are less likely to make generic substitution, and that the describing doctor is less likely to accept generic substitution if they are males, in particular when they are getting older. Thus we assume:

(12) $$v_{nGt} = \alpha_G + \beta_1 P_{nGt}$$

(13) $$v_{nBt} = \alpha_B + \beta_1 P_{nBt} + \beta_2 \text{Patientage}_{nt} \text{ x Male}_n + \beta_3 \text{Doctorage}_{nt} \text{ x Male}_n$$

where

$P_{nGt} = \text{price of generic}$
$P_{nBt} = \text{price of brand}$

The prices may vary across time and patients. It should be noted, however, that for all individuals social security cover part of the expenses on *statines*. This is accounted for in the paper.

From the structure of the model we easily see that we can only identify $\alpha_B - \alpha_G = \alpha$. Our expectation with respect to the sign of the coefficients are $\beta_1<0$, $\beta_2>0$, $\beta_3>0$.

The model implies the following price-elasticities:

(14)

$$
\begin{aligned}
&(a)\quad ElQ_{nBGt}:P_{nGt} = \beta_1 P_{nGt} Q_{nBBt};\ \text{for } t>t_0\\
&(b)\quad ElQ_{nBBt}:P_{nGt} = -\beta_1 P_{nGt} Q_{nBGt};\ \text{for } t>t_0\\
&(c)\quad ElQ_{nBGt}:P_{nBt} = -\beta_1 P_{nBt} Q_{nGBt};\ \text{for } t>t_0\\
&(d)\quad ElQ_{nBBt}:P_{nBt} = \beta_1 P_{nBt} \frac{Q_{nGBt} Q_{nBGt}}{Q_{BBnt}};\ \text{for } t>t_0\\
&(e)\quad ElQ_{nGBt}:P_{nGt} = -\beta_1 P_{nGt} Q_{nBGt};\ \text{for } t>t_0\\
&(f)\quad ElQ_{nGGt}:P_{nGt} = \beta_1 P_{nGt} \frac{Q_{nGBt} Q_{nBGt}}{Q_{nGGt}};\ \text{for } t>t_0\\
&(g)\quad ElQ_{nGBt}:P_{nBt} = \beta_1 P_{nBt} Q_{nGGt};\ \text{for } t>t_0\\
&(h)\quad ElQ_{nGGt}:P_{nGt} = -\beta_1 P_{nBt} Q_{nGBt};\ \text{for } t>t_0
\end{aligned}
$$

3. The data

Our data were extracted from the Norwegian Prescription Database (NorPD) at the Norwegian Institute of Public Health. The NorPD (Norwegian title: Reseptregisteret) was established on 1st January 2004.[2] The Database monitors all drugs that are dispensed by prescription in Norway, and provides information about the patient (age, sex, and insurance status), the physician (age, sex, and speciality), the pharmacy (location), and the dispensed drug (price, package size, strength, product name). Using other sources of information provided by the Norwegian Medicines Control Authority (list of pharmacies and a list of drugs approved for the Norwegian market), we get additional information about pharmacy ownership, identity of the main wholesaler and producer name and price of the drugs. The latter is used to identify brand-name drugs and generics.

In the data set only the price of the drug chosen (p_dd*)* is reported that may be brand or generic. To generate the price of the drug not chosen (p_not) we have done as follows.
First we generated a dummy variable (*b_chosen*) that identify if the drug is brand or generic. It is equal to one if the drug name is Pravachol or Zocor (alone or in combination); atc_code is C10AA03 or C10AA001, 0 otherwise. Then, we generated the

[2] See Furu (2001)

mean price (*p_ddd*) over the chosen drug that has same *atc_code*, same strength (*strength)*, same pharmacy identifier (*id_n_ph)* and same date of transaction (*months*). At last we generated the alternative price (*p_not*) equal to the mean price just computed, conditioned on *b* (1 or 0). It happens that there are groups in which only brand is chosen or only generic is chosen. In these cases we could not compute the alternative price and we then set *p_no*t equal to missing. It also happens that in some groups there is just only one observation useful to compute the average. Also in this case we set the value of *p_not* to missing. To sum up:

p_generic = p_ddd*(1-b_choicen) + p_not*b_choicen;

p_brand = p_ddd*b_choicen + p_not*(1-b_choicen);

where: *p_generic* is the price of the generic drug; *p_brand* is the price of the brand drug, *p_ddd* is the price of the chosen drug and *p_no*t is the price of the drug not chosen, and *b_choice* is a dummy variable equal to 1 if brand is chosen and 0 otherwise.

In the sample there are at least 28 prescriptions by patients over the 37 months. We observe drug prescriptions from May 2004 (first prescription considered) to June 2007 (month 5 to 42), a total of 37 months.

After the selections listed above, we get 3898 observations that refer to 109 patients.
The panel is unbalanced since for each patient there are a different number of prescriptions from May 2004 to June 2007.

The following statistics, show that at minimum a patient has 28 prescriptions, and at maximum 52 prescriptions.

Table 1. Number of prescriptions, May 2004-June 2007, statines

Number of patients	Mean no of prescriptions	Std.Dev	Min	Max
109	35,76	6.33	28	52

The number of prescriptions by patient is not equal to the number of months since there may be more than one prescription per month.

Table 2 gives the description of the variable while Table 3 gives the descriptive statistics.

Table 2 Description of the variables

Variable	Description
p_ddd	price (in NOK) per daily dose (i.e. p_ddd = no_packages * p_packages/no_ddd)
p_not	price of not chosen
b	Dummy: b = 1 if brand (drug_name is equal to "Pravachol" and atc_code is equal to "C10AA03" or drug_name is equal to "Zocor" and atc_code is equal to "C10AA01"), b = 0 if generic (i.e. Pravastatin and Simvastatin)
p_generic	Price per daily dose of generic drug
p_brand	Price per daily dose of brand drug
age_d	Age of the doctor
age_p	Age of the patient
patient_m	Dummy: 1 if male, 0 otherwise
patient_f	Dummy: 1 if female, 0 otherwise
doctor_m	Dummy: 1 if male, 0 otherwise
doctor_f	Dummy: 1 if female, 0 otherwise
months	months of drug prescription ranges from 5 (May 2004) to 42 (June 2007)

Table 3. Descriptive Statistics (number of observations 3898 – 109 patients)

Variable	Mean	Std.Dev.	Min	Max
p_ddd	2.7184	1.9349	0.5679	9.7388
p_not	3.6193	2.1647	0.8693	9.6857
b	0.1637	0.3700	0.0000	1.0000
p_generic	2.3705	1.1450	0.5679	7.0850
p_brand	3.9671	2.5000	0.8694	9.7388
age_d	50.5872	9.3071	29.0000	68.0000
age_p	78.4254	8.6641	50.0000	91.0000
doctor_f	0.1329	0.3395	0.0000	1.0000
doctor_m	0.8671	0.3395	0.0000	1.0000
patient_f	0.4115	0.4922	0.0000	1.0000
patient_m	0.5885	0.4922	0.0000	1.0000
months	28.9690	8.6594	5.0000	42.0000

4. Results

Table 4 gives the estimates. We observe that price has the expected negative impact on demand and the impact is significant different from zero. The interaction of male doctors and age has a positive and significant impact on the use of brand products. Patient's age interacted with gender has no significant impact

The preference discount factor is positive and significant which indicates that preferences are correlated over time, given the covariates in the deterministic part of the utility function.

Table 4. Estimates.

Variables	Parameters	Estimates	t-values
Constant	α	3.2152	15.337
Price	β_1	-1.1913	-2.841
Patient age x Male	β_2	-0.0373	-0.965
Doctor age x Male	β_3	0.2096	3.967
Preference discount factor	θ	3.7475	4.249
No of observations	3898 (109 patients)		
Mean log-likelihood	- 433.126		

Correlation matrix of the estimated parameters

	α	β_1	β_2	β_3	θ
α	1.000	-0.454	-0.003	-0.118	0.572
β_1	-0.454	1.000	-0.279	-0.228	0.123
β_2	-0.003	-0.279	1.000	-0.352	-0.048
β_3	-0.118	-0.228	-0.352	1.000	-0.029
θ	0.572	0.123	-0.048	-0.029	1.000

From Table 5 we observe that all elasticites have the expected sign, which of course come the fact that $\beta_1<0$. The only two sizeable elasticities are the most important ones. The elasticity of transiting from brand to generics (statines) with respect to the generic price is on average equal to -0.2732**.** The elasticity of transiting from generics to brand (statines) with respect to the brand price is on average equal to -0.4625. The brand price has thus a stronger impact on the the transition than the generic price. In Figur 1 we show how the elasticities vary across the 37 months. We observe that the two most important elesticities referred to above indicate that price responses were strongest at the beginning of the period (May 2004) and at around month 20 (January 2006)

Table 5. Elastisites of the transition probabilites with respect to prices; averaged over patients and periods.

a) for transition from brand to generic as a consequence of an increase in generic drug price (see eq14 a)
b) from brand to brand as a consequence of an increase in generic drug price (see eq. 14 b)
c) for transition from brand to generic as a consequence of an increase in brand drug price (see eq. 14 c)
d) from brand to brand as a consequence of an increase in brand drug price (see eq. 14 d)
e) for transition from generic to brand as a consequence of an increase in generic drug price (see eq 14 e)
f) from generic to generic as a consequence of an increase in generic drug price (see eq. 14 f)
g) for transition from generic to brand as a consequence of an increase in brand drug price (see eq. 14 g)
h) from generic to generic as a consequence of an increase in brand drug price increase (see eq. 14 h)

Variables	Mean	Min	Max	Std.
eq. 14 a) ElQ_{BGnt}:_P_G	-0.2732	-0.8416	-0.0658	0.1293
eq.14 b) ElQ_{BBnt}:_ P_G	0.0092	0.0006	0.1189	0.0111
eq 14 c) ElQ_{BGnt}:_P_B	0.0101	0.0006	0.1405	0.0142
eq 14 d) ElQ_{BBnt}:_P_B	-0.0003	-0.0008	-0.0001	0.0002
eq 14 e) ElQ_{GBnt}:_P_G	0.0092	0.0006	0.1189	0.0111
eq 14 f) ElQ_{GGnt}:_P_G	-0.0002	-0.0006	0.0000	0.0001
eq 14 g) ElQ_{GBnt}:_P_B	-0.4625	-1.1406	-0.1011	0.2878
eq 14 h) ElQ_{GGnt}:_P_B	0.0101	0.0006	0.1405	0.0142

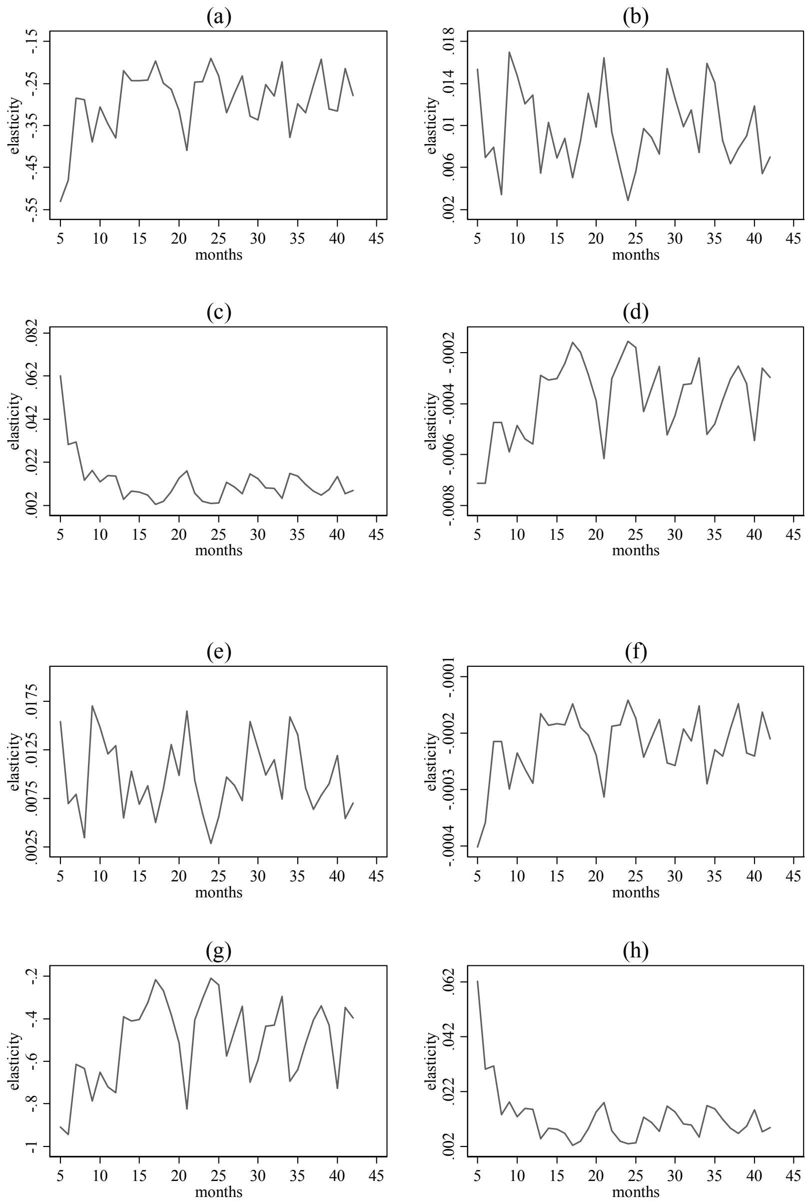

Figure 1.Mean elasticity of probability vs. month

a) for transition from brand to generic as a consequence of an increase in generic drug price (see eq. 14a)
b) from brand to brand as a consequence of an increase in generic drug price (see eq. 14 b)
c) for transition from brand to generic as a consequence of an increase in brand drug price (see eq. 14 c)
d) from brand to brand as a consequence of an increase in brand drug price (see eq. 14 d)
e) for transition from generic to brand as a consequence of an increase in generic drug price (see eq. 14e)
f) from generic to generic as a consequence of an increase in generic drug price (see eq. 14 f)
g) for transition from generic to brand as a consequence of an increase in brand drug price (see eq. 14 g)
h) from generic to generic as a consequence of an increase in brand drug price increase (see eq. 14 h)

In Table 6 we report the mean of the correlation of utilites across patients (and time). When the drug type is the same, the correlation is mainly due to the coefficient θ, the preference discount factor. When the drug types are different (B and G) the correlation is also affected by the fact that the characteristics of the different drug types differ. Figure 2 gives the variation across all 109 patients. Table 7 report the same correlation across time and Figure 3 show how these correlations varied over the 37 months.

Table 6. Mean correlation of utilities for the 109 patients.

$\text{corr}(\exp(-U_{nj}(t-1)), \exp(-U_{ni}(t))$	Mean	Std. Dev.	Min	Max
j=B, i=B	0.023352	0.0007092	0.0222853	0.0260270
j=G,, i=G	0.023417	0.0007232	0.0219913	0.0260024
j=B, i=G_	0.277621	0.1257965	0.1303722	0.8622116

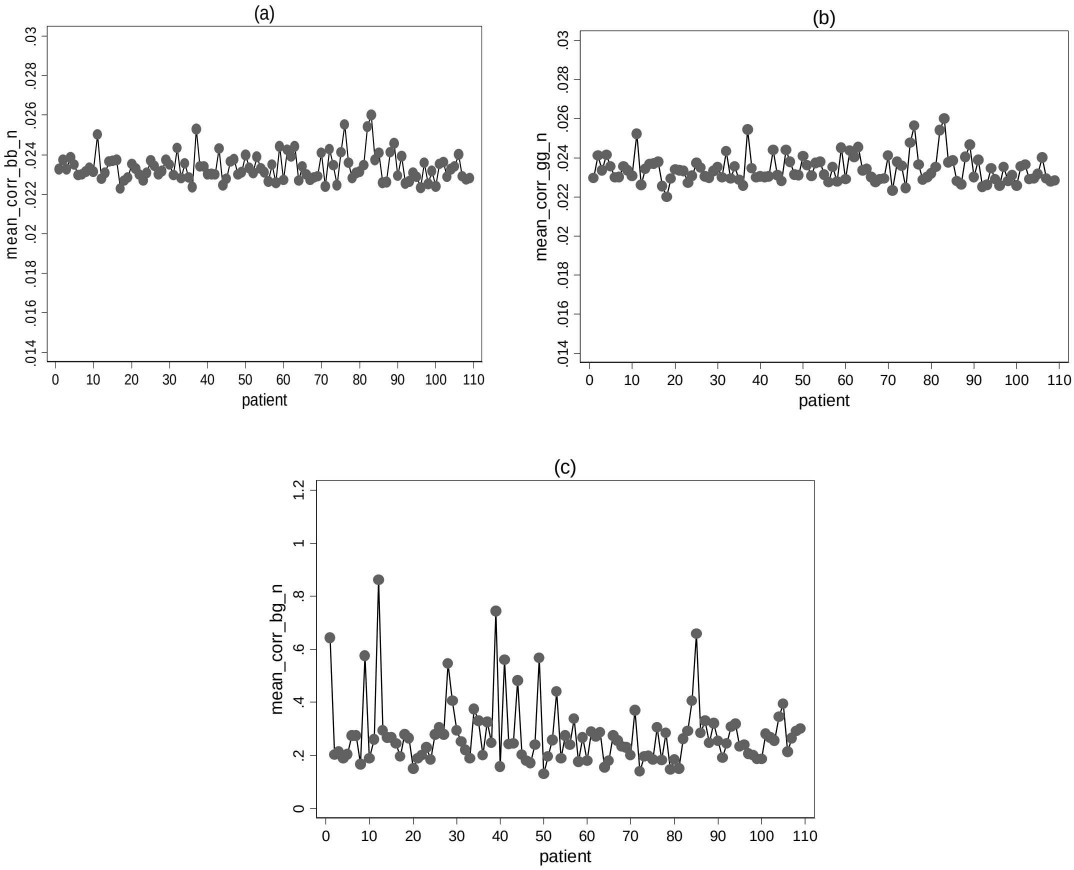

Figure 2. Mean correlation of utilities: (a) mean correlation of transition from brand to brand (b) mean correlation of transition from generic to generic, (c) mean correlation of transition from brand to generic

Table 7. Mean correlation of utilities across time

$\mathrm{corr}(\exp(-U_{nj}(t-1)), \exp(-U_{ni}(t)))$	Mean	Std. Dev.	Min	Max
j=B, i=B	0.0218287	0.0034909	0.0117883	0.0252339
j=G,, i=G	0.0219159	0.0033689	0.0117883	0.0253417
j=B, i=G_	0.2536082	0.0523233	0.094179	0.3145396

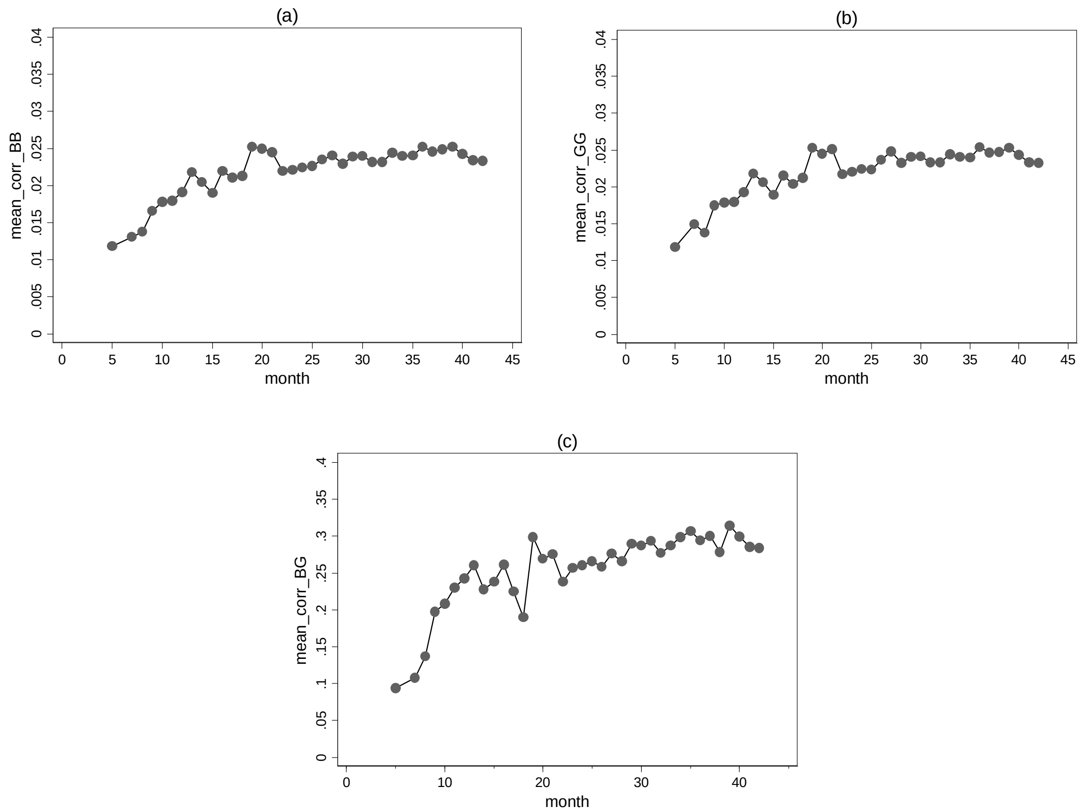

Figure 3. Mean correlation of utilities across time: (a) mean correlation of transition from brand to brand (b) mean correlation of transition from generic to generic, (c) mean correlation of transition from brand to generic.

5. Conclusions

Using an extensive longitudinal dataset extracted from the Norwegian Prescription Database (NorPD) containing all prescriptions written in the period January 2004 to June 2007, we selected two particular drugs (chemical substances) used against cholesterol. The two brand-name products on the Norwegian markets were *Provachol* (atc code C10AA03) and *Zocor* (atc code C10AA01). The generics are *Provastatine* and *Simastatine*. The model accounts for taste persistence and is estimated on panel data. We find that prices have a negative impact on transitions in the sense that an increase in the brand price will reduce the transition from generics to brand and likewise an increase in the generic price will reduce the transition from brand to generics.

References

Dagsvik, J.K.(2002): Discrete choice in continuous time: Implications of an intertemporal version of the IIA property. *Econometrica,* 2002, **70**, 817-831

Dalen, D. M., M. Locatelli and S. S. Strøm (2010): Generic substitution: Micro evidence from register data in Norway, forthcoming in *European Journal of Health Economics*

Furu, K. (2001), Drug utilisation in a public health perspective: Establishing a national prescription register in Norway. *Norwegian Journal of Epidemiology*, 11 (1): 55-60.