Learning by Doing and Consumer Switching Cost

Yufeng Huang[†]

April 19, 2013

Preliminary. Please to not cite or distribute without contacting the author.

Abstract

What internal factors can induce consumer switching, even when the environment is stable? In this paper, I characterize the relationship between the endogenous evolution of consumers' product-specific human capital (in this case, their capability in producing good pictures using a digital camera), and their decisions in camera purchasing and picture production. Simply put, the more experience a consumer has in producing pictures, the better picture quality she could produce (learning by doing), and hence a higher willingness to transition into better cameras. In addition, I find that despite single lens reflex (SLR) cameras have much higher potential to generate high quality pictures, they are usually much more difficult to handle, especially for a consumer who is used to a simple compact camera. Such lack of user-friendliness in advanced products generates a switching cost.

If firms are willing to improve their product user-friendliness – for products such as the SLRs with high potential but steep learning curve – there is room for welfare gains. I estimate the consumer willingness to pay for having their switching cost removed to be as high as 26% equivalent of the price of SLRs. On the other hand, a manufacturer who undertakes product-redesign efforts to remove the switching cost, is capable of charging as high as an 18% SLR price premium. This shows that it is both socially and managerially desired to decreased consumer switching cost – despite some incentive misallignment.

Keywords: switching cost, learning by doing, dynamic programming, forward-looking behavior, digital camera market

---

I am grateful to my advisers – Bart Bronnenberg and Tobias Klein, for their constant guidance and support. This paper also benefit from helpful comments and suggestions from Jaap Abbring, Bart Bronnenberg, Tobias Klein, Carlos Santos and Benjamin Tanz, and seminar participants in Tilburg and Mannheim. I thank The Netherlands Foundation for Scientific Research (NWO Vici Grant 453-09-004) for financial support. All errors in this paper are mine.

[†]Tilburg University, Email: `y.huang@uvt.nl`.

1 Introduction

When the environment (objective product characteristics and budget constraints) is stable, what are the internal factors that could induce consumers to switch among products? To name two examples. In Israel (2005), drivers tend to switch car insurance providers after incidence of a claim, because of dissapointment of the service, or *familiarity*; while in our case, consumers who are better at taking pictures, tend to transition to a single lens reflex (SLR) camera, which is potentially more capable of producing high quality pictures. The second is informally known as "*expertise*".

In the traditional marketing language such as in Alba and Hutchinson (1987), these two examples are two explanations to internal factors that generate nonstationary consumer decisions in a stationary environment. While the first – familiarity – has been formally discussed by the huge literature on uncertainty and information,[1] few have characterized the role of "expertise" in a consumer decision context. However, the notion of how consumers become better at interacting with products, could be key to explain why they tend to move up in the product quality ladder – a pattern usually seen in sports equipments, consumer electronics, and many other categories.

This paper studies consumers' incentives in switching among digital cameras. I characterize consumers' increasing tendency to choose cameras of higher quality, in relation to the endogenous evolution of their product-speicifc *human capital* (Schultz, 1961; Arrow, 1962; Becker, 1965). [2] Simply put, a product's physical quality and the consumer's product-specific human capital ("human capital" for abbreviation) are complementary in determining the efficiency in interacting with the product, and ultimately the consumption utility. This suits well in the camera-purchase context, as producing high quality pictures is one of the ultimate goals of owning a camera.[3]

It is crucial to point out that human capital is not a permanent individual characteristics, and in addition, its evolution is endogenous to one's choice history. On the one hand, human capital accumulates as the individual develops better understanding of the camera she owns, by taking more pictures with it – a classical mechanism known as *learning by doing* (Arrow, 1962). On the other hand, different cameras might be drastically different in their design, functions or hidden caveats, so that the knowledge developed

[1] The Bayesian learning literature such as Erdem and Keane (1996); Crawford and Shum (2005); Osborne (2007), or the search literature as in Moshkin and Shachar (2002), among others.

[2] I do not use the general term "expertise" since *human capital* is very specific to a content production context, which will be central to this paper.

[3] Cf. consumer surveys in CEA (2012).

in one camera might not be fully transferable to the other, should the individual decides to switch among products. For example, to an individual who is used to simple entry level cameras, a more advanced camera might be much more difficult to operate well initially – hence generating a switching cost in the human capital which slows down the individual from moving up in the camera quality ladder.

The theory is tested on a unique panel data I collected from Flickr.com. The data is unique because it is detailed on the characteristics of pictures the individual produces – from which I infer their picture quality in each period. In light of the reduced-form evidence, I construct a dynamic discrete choice model that jointly captures the evolution of individual human capital, her incentives to take a picture (which is both production of a consumption good and an investment in her human capital), and her incentive in switching among cameras. The model provides intuitive patterns in price responses, the trade-off between higher potential picture quality and higher-priced cameras, and evolution of the distribution of human capital. It also fits the evolution of picture quality and dispersion of single lens reflex (SLR) cameras very well.

The parameter estimates suggest strong evidence for consumer switching cost from moving up the quality ladder: due to the fact that SLR cameras are more complicated to use, switching from a compact camera to an SLR generates an attrition in the consumer human capital, to such an extent that for some consumers, it even more than offsets the immediate incremental camera quality. This seems to suggest that there is potential improvemet in welfare, from improving the user-friendliness of SLR cameras. In the counterfactual experiments, I quantify the welfare benefits from re-designing all SLR cameras, so that a consumer transitioning to one of these cameras no longer loses any human capital. My counterfactual experiment shows that, the consumers are willing to pay for, on average an 18%-26% premium in the SLR prices,[4] for a policy that eliminates all switching cost to SLRs. On the other hand, a firm is only capable of exploiting 10%-14% in the SLR prices, keeping consumers at equal switching probability to SLR cameras. In a sense, this is the benefit – to a society and to a firm, respectively – from designing SLRs that are “as user-friendly as” the simple compact cameras. Due to the high consumer surplus which cannot be captured by firms, it is not surprising to see a lack of incentives from the supply side to undertake such a product re-design task.

This paper is closely related to the theoretical and empirical literature on Bayesian learning by doing and product adoption. The seminal theoretical framework by Jovanovic and Nyarko (1996) serves as our

[4]Depending on which set of parameter estimates I use to generate the counterfactuals.

back-bone for model bulding, despite that I model a dynamic decision problem while theirs is static. In the empirical literature, a similar framework is applied to study farmer's decision of adopting agricultural technologies in Foster and Rosenzweig (1995), who examine the decision of adopting a new technology and the land productivity for Indian farmers. The central difference in this paper, is that choice of SLRs is not an absorbing state as one can always switch back and forth – a crucial variation in identifying consumer switching cost from the effect of having a higher human capital.[5]

To my knowledge in the marketing and empirical industrial organization literature, this is the first paper to structurally examine, for a consumer with rational expectations, endogenous evolution of product-usage efficiencies (human capital) and its impact on the incentives to switching among products.[6] However, we fit in (or are similar to) a broader literature on dynamic models that characterize endogenous evolution in consumer tastes – in particular the ones on uncertainty and (Bayesian) learning. For example, Erdem and Keane (1996) is one of the early works that attributes consumers' tendency to purchase products that i) they have purchased before or ii) they were informed of in the advertisements, to the role of uncertainty in product quality, risk aversion in preference, and Bayesian learning. Works that are similar in essence include Crawford and Shum (2005); Israel (2005); Osborne (2007). This paper presents a different mechanism from the Bayesian learning literature: here, consumers' tendency to adopt new cameras is not because of familiarity, but their capability to use the new cameras well. We have different welfare and managerial implications.

The remainder of the paper is organized as follows. Section 2 discusses how the data is collected, while some important descriptive evidence – which justifies a model – will be presented in Section 3. Next, Section 4 presents the model and discusses how it can be estimated. Results and counterfactual experiments are presented in Sections5 and 6, respectively. Finally, Section 7 concludes.

[5]In Foster and Rosenzweig (1995), decision to using the new technology is absorbing, in the sense that no farmers ever switches back in their data.

[6]For potentially related laboratory evidence, cf. Alba and Hutchinson (1987).

2 Data

2.1 Consumers In The Digital Camera Industry

The digital camera industry is a large sub-sector in consumer electronics. *Excluding* camcorders, cell-phones, tablets or any other multifunctional devices, wholesaling in digital cameras in the US amounts to 7 billion, which takes up 0.85% of the entire US wholesale GDP.[7] On the other hand, the market is fairly oligopolistic, with 61% of the market share in digital cameras taken by the top 4 manufacturers: Canon, Sony, Nikon and Samsung.

On the consumer side, an interesting phenomenon is that consumers purchase a new camera before their current camera wears out. According to a survey by the Consumer Electronics Association (CEA), only 26% of the respondents report that they purchase when the current device fails to work; contrasting the 55% of consumers who wants a higher-quality device or "latest technology".[8] Other similar patterns have been found by the field surveyors; for example, consumers are most likely to be dis-satisfied with their current camera because of "poor picture quality".

2.2 Data Collection

I extracted my data from the *Flickr.com*, using a set of self-coded programs. Launched in February 2004 by Ludicorp, Flickr was acquired by Yahoo! in March 2005 and merged with the existing Yahoo! Photos (which was originally launched in 2000). By Summer 2012, it was ranked #52 globally in internet traffic.[9]

The advantage of data availability comes from the publicity nature of Flickr as an image hosting website, where most pictures can be seen by an anonymous visitor. Flickr is one of the most popular image hosting websites. My primary data comes in two levels: picture level and individual level. On the picture level, one is able to observe the *exchangeable image file format* (Exif) data recorded by the camera when a picture was taken. This data contains information of the time and date when a photo was taken, and the equipment that does the job. In addition, it contains detailed camera settings such as shutter speed, aperture and exposure mode. By default, such data is displayed on Flickr as public information.[10] Besides the Exif data, Flickr records the time and date when a picture was uploaded, as well as its *cumulative*

[7] Reports from Consumer Electronics Association Sales and Forecast, January 2012. See www.ce.org.
[8] CEA (2012), via www.pmai.org.
[9] Alexa.com, August 11, 2012.
[10] And more than 70% of the pictures in our sample contain this information.

number of views, comments and "favorites" since upload.[11] Besides picture level data, I also extracted a cross-section of individual (user) level data, which directly provides a summary of the number of pictures taken and uploaded in each month.

In additional to the data above, I collect a cross-sectional data-set for camera characteristics, and a longitudinal price data. To do this, I combine information from different sources: Flickr provides characteristics such as resolution (mega-pixels) and weight, for a total of 2,408 cameras.[12] Dpreview.com lists camera characteristics as well as launch dates for 1,913 cameras – so that we can infer camera availability. Cnet.com provides information for 1,070 cameras on their manufacturer suggested retail price (MSRP). And finally, Pixel-peeper.com lists longitudinal data as long as 6 years, on average Ebay transaction prices for 767 cameras.

2.3 Sample Selection and Summary Statistics

To create a sample of pictures that can well represent the camera user's ability, I focus only on the Pro accounts (paid accounts).[13] This is because most of the free account users tend to upload very few pictures throughout their entire course of usage. Not only would one have inaccurate inference regarding the evolution of their human capital over time (because of the limited number of observations), one might also run into the risk of omitting certain camera-switching decisions because of lack of observation on their pictures. On the other hand, however, Pro accounts are less representative for the entire market. This is arguably less of a concern for this paper, since my primary focus is on within-individual changes rather than across-individual differences. However, the quantitative conclusions from the counterfactual experiments of this paper cannot be over-generalized.

Focusing on Pro accounts gives a sample size of 7077 users (Table 1), from a random sample of user names. For these users, I collect relevant information from the first in every five pictures that they have uploaded.[14] In total, 6,781 individuals have multiple observations on their picture quality and camera they

[11] A "view" is a click on a thumbnail of a picture, so that the viewer can look at an enlarged version of it. Multiple clicks by a viewer is counted as one. A registered viewer can leave a comment to a picture. A viewer can also choose to indicate that the picture is her "favorite" – a feature similar to the Facebook "like". Yet, most pictures have not received any favorite votes, making it difficult to directly use this variable as indication of picture quality.

[12] Part of this information is missing, especially for non-popular cameras.

[13] For registered users, Flickr offers two types of accounts: i) a free account, which is imposed a monthly uploading capacity (300 megabytes as in 2012), as well as the restriction that only the most recent 200 photos are observable (even to the user himself); and ii) a Pro account, requiring an annual subscription fee ($24.95 in 2012), but is alleviated from all the above restrictions.

[14] More accurately, I visit the first of every 5 pictures on *each page* in the user's photo-stream. Collecting picture level

Table 1: User Level Data Summary

	Mean	Stdev
	b	b
total number of pictures	1878.988	2115.188
number of in-sample pictures	376.878	461.159
number of votes to others' photo	276.819	1786.164
number of contacts	94.693	348.609
obs.	7077	7077

Note: users with no more than 200 total uploads are excluded.

Table 2: Picture Level Data Summary

	All		With Complete Exif	
	Mean	StDev	Mean	StDev
	b	b	b	b
views	55.634	1223.723	47.164	1233.497
comments	0.585	5.778	0.530	5.821
favorites	0.368	7.258	0.316	7.317
shot by a cellphone	0.038	0.192	0.044	0.192
shot by a compact camera	0.385	0.487	0.445	0.487
shot by an SLR	0.404	0.491	0.465	0.491
shot by a mirrorless camera	0.010	0.100	0.011	0.100
shot by other cameras	0.030	0.172	0.035	0.172
Exif data incomplete	0.131	0.338		
months in display			34.404	22.273
months since first picture taken			54.812	33.336
obs.	2667169	2667169	2287333	2287333

use for taking those pictures, and the resulting picture-level data has close to 3 million observations (see Table 2). Among this data, we delete pictures with incomplete Exif data (13%),[15] and pictures taken by cellphones (4%), mirrorless cameras (1%) and camcorders or film cameras (3%) – leaving us with 79% usable data.

information for all of them will greatly enlarge the data-set and thus the time required to extract it, yet might provide only minor improvement in the standard errors.

[15]Considered there are pictures which do not contain an Exif data *per se*, I do not consider the case where individuals selectively delete Exif data before uploading.

2.4 Proxy of Picture Quality

Ideally, one would use average ratings as a proxy to picture quality. Nevertheless, a majority of the pictures in our sample has zero “favorite” votes (see Table 2), making this variable unusable for our purpose. Instead, I develop a proxy for picture quality from the viewer-count data, i.e. the cumulative number of views for each picture until the time of data extraction. Intuitively, views could reflect picture quality since viewing a picture takes time,[16] thus a viewer is likely to select the highest quality picture from a set of thumbnails she observes.

The primary problem is to isolate quality from popularity, since the number of views might reflect high popularity of the photographer or the subject, rather than high quality of the picture.[17] To do this, I exploit variation in the date of picture-taking and the date of picture-posting, among other variables, to identify the quality of the picture – realized when the picture was taken – from popularity factors, which is realized after the picture posted.

A simple example could illustrate this. Suppose two pictures from *a given photographer*, A and B, are both taken at t_0 but posted at t_A and t_B respectively. Then, it is plausible to assume that picture A and B have no systematic difference in their quality, and any difference in view counts is attributed to popularity effects. On the other hand, if another pair of pictures, a and b, are taken at t_a and t_b respectively but posted on the same day t_1, then I assume no difference in the popularity effects and their difference in view counts attributed to difference in quality.

To implement this idea, I regress the log cumulative number of views for each picture on a set of control variables that captures the popularity effects, and then take the average residual and individual fixed effects as proxied picture quality. The specification of regressors include: i) a polynomial of time dummies that a picture is in display, ii) number of pictures taken in the same month; iii) upload activities of the individual when the picture is in display, and iv) the number of views of the 20 adjacent pictures. Individual and camera fixed effects are controlled. Here, a group of display time indicators capture the popularity effects through variations in the date of picture-posting, upload activities captures backward spillovers from popularity of the pictures that are posted later (in spirit of Hendricks and Sorensen, 2009),

[16]By accounts from Alexa.com, “Visitors to the site view an average of 9.8 unique pages per day. Visitors to Flickr spend about five minutes per visit to the site and 25 seconds per page-view”. See http://www.alexa.com/siteinfo/Flickr.com.

[17]Or put it differently, in a counterfactual world when people are to choose which picture to view after observing *all* thumbnails on Flickr, they would not have chosen some of the pictures that, in this world, have the highest view counts.

Table 3: Compare with Alternative Quality Measures

	overall	positive favorites/comments
mean	-0.018	2.094
overall s.d.	1.404	1.234
within-i s.d.	0.912	0.857
within-i-cam s.d.	0.910	0.930
within-i-cam-t s.d.	0.742	0.813
corr. with log favs		0.525
corr. with log comments		0.450
corr. with favs net of popularity		0.516
obs.	2584750	130200

Notes: The second to fourth rows present standard deviations of the predicted picture quality, for the entire sample, within individuals, within individual-camera combinations and within individual-month-camera combinations, respectively.

and so does the popularity of the adjacent pictures. Details of this specification is presented in Section B in the Appendix.

To further examine whether the residuals can proxy picture quality, I compared the residuals to (log) number of favorite votes and comments for a subset of pictures, where their votes and comments are nonzero. This result is shown in Table 3, together with some summary statistics of this variable. Remarkably, the correlation coefficient to both (log) number of comments and favorites are close to 0.5.

3 Reduced-Form Evidence

Before imposing any structural modeling prior, I will first present some simple reduced-form evidence to justify the use of my model. First and foremost, there is a discontinuity in the cumulative number of views, when an individual switches from a compact camera to an SLR. Given that viewers do not necessarily know that camera switching occurred, this suggests a switching cost that negatively affects picture quality.[18] Secondly, there is a upward-sloping trend in the quality of pictures one produces –

[18] The same pattern occurs if one replaces cumulative number of views by the proxied picture quality.

indicated by a trend in the views even when numerous popularity effects are controlled. Finally, I will show that the decisions to take pictures is endogenous to the potential picture quality one could produce.

3.1 Switching Cost in Human Capital

This section first presents model-free evidence on switching cost in consumer human capital – which is the central subject of this paper. As figure 1 shows, at the point when the consumer switches camera *type*, the pictures she produces enjoy less views than otherwise – even when these pictures were uploaded at a later period. This is true even when the consumer switches to an SLR, in which case the camera will produce pictures that have much higher views later on. These patterns suggests that the differences in views generated by different cameras are not purely due to the cameras' physical quality, and in addition, the user's capability in using the camera (human capital) is subject to a negative impact at the point of switching.

One alternative hypothesis is that the discontinuity might be due to differences in popularity of the camera. However, if this is true, then discontinuity in views should show different signs at the left and right panels of Figure 1. In particular, when a consumer switches to an SLR which later generates higher popularity, then it should generate higher popularity at the point of camera switching.[19] This pattern is also robust when one instead examines (proxied) picture quality, which eliminates popularity effects and changes in selection criteria.

3.2 Learning by Doing

Figure (1) also shows an upward-sloping trend in the number of views, in the first few years after switching to a new camera. To examine whether such time trend exists in general, I use a standard reduced-form specification to testing learning by doing, as in Sheshinski (1967). Namely, we test for the hypothesis that (log) number of views is increasing in the time periods since a user starts to take pictures, controlling for popularity effects (as in our ways to proxy picture quality). Further, we examine whether the number of periods when the individual produces a picture explains the time trend.

The first column of Table 4 shows a positive and concave time trend in the number of views, which indicates that, after controlling for factors that reflect the evolution of popularity of a photographer, pictures

[19] In fact, since Exif data is not shown on the page of the picture, many viewers might not even notice the camera switch if the user does not explicitly mention it.

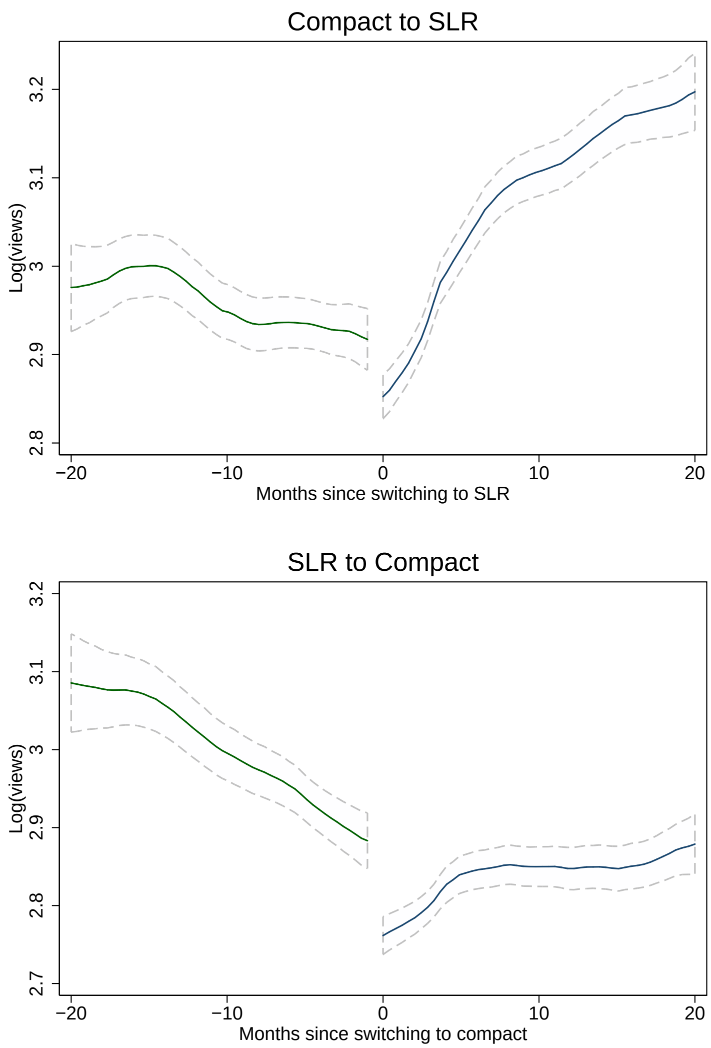

Figure 1: Log Cumulative Number of Views and Months after Switching

Notes: The variable on the vertical axis is the proxied picture quality, net of all camera dummies. The detailed specification is documented in Appendix B.2.

Table 4: Tenure and the Number of Views

	$\log(Views/Months)$	$\log(Views/Months)$
Months in Flickr (00s)	0.671***	-0.008
	(0.010)	(0.013)
Months squared (0000s)	-0.159***	0.127***
	(0.007)	(0.008)
Months spent taking pictures (00s)		1.524***
		(0.021)
Months taking pictures squared (0000s)		-1.356***
		(0.017)
Display month dummies	Yes	Yes
Views of adjacent pics	Yes	Yes
Pictures taken now	Yes	Yes
Pictures uploaded now and subsequently	Yes	Yes
Individual and camera fixed effects	Yes	Yes
Rsq.	0.121	0.124
obs.	2667169	2667169

Note: Within individual R-squared are provided.

taken later enjoy higher number of views. In addition, as shown in the second column, the number of time periods when the user has taken at least one picture ("months spent taking pictures") and its quadratic term, seem to explain most of the time trend found in the first column. This strongly supports the learning by doing hypothesis (Arrow, 1962; Sheshinski, 1967) – in our context, the photographic specific human capital evolves ("learn") as one takes more pictures ("do"). Quantitatively, a month of picture taking boosts the quality of one's pictures such that they attract 1.5% more viewers.

3.3 Picture-Taking Activities

Before constructing a model of human capital accumulation, it is also crucial to know whether evolution of human capital through picture-taking is endogenous. One hypothesis is that individuals need to produce pictures in order to consume them – so that the incentive to take pictures is endogenous to the potential picture quality one might be able to produce. Alternatively, from an endogenous investment perspective, picture taking might be endogenous to the marginal return of investment – which is decreasing to the level of human capital.

To provide some evidence evidence, I plot the probability of taking at least one picture, against the (linearly) interpolated picture quality for all individuals. Interpolated picture quality can be seen as a

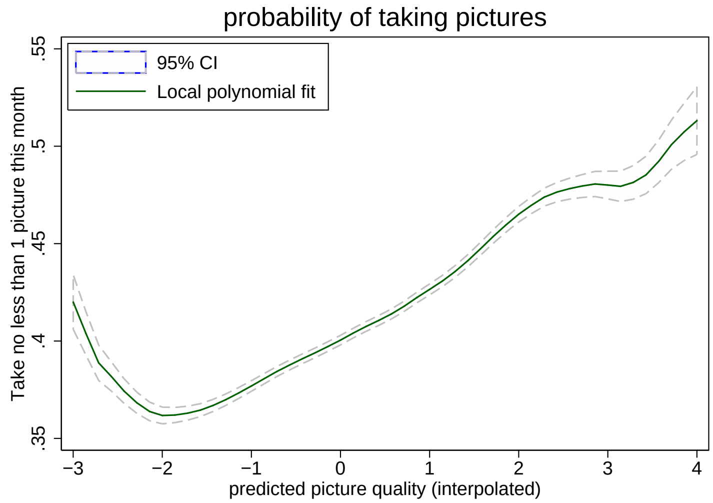

Figure 2: Probability of Picture Taking and Interpolated Picture Quality

Notes:

crude representation of the potential picture quality one would have produced in a given period. Figure 2 shows evidence of both an investment incentive an a consumption incentive to taking pictures – the latter dominates the former when potential picture quality is larger. This indicates that i) we need to model the decision of taking pictures, and ii) picture taking is both an investment in human capital and production of a consumption good.

4 Model and Econometrics

4.1 Overview

As shown in the reduced form evidence, a structural model needs to characterize consumers' purchase decisions in response to i) the switching cost in their picture quality, and ii) the increasing and concave trend of their human capital, especially with respect to the number of picture-taking periods. In addition, the picture taking decisions need to be endogenized to allow for the incentive to do so, to vary across different human capital levels.

In this section, I present a model to jointly characterize consumers' purchase decisions of cameras, and their decisions to take pictures – in relation to the evolution of their human capital. In the model, production of a picture generates, on the one hand a consumption good that brings utility to the consumer, and

on the other hand, a signal of how a picture can be correctly taken, which contributes to the improvement of consumer human capital – through a Bayesian learning by doing structure *a la* Jovanovic and Nyarko (1996). Having rational expectations, the consumer decides whether it is worth spending the effort to take a picture in each period, and whether it is worth spending the income in purchasing a (potentially different) camera.

4.2 Decision and Timing

The central problem is a dynamic decision process where a consumer makes decisions on purchasing durable goods, i.e. cameras. Consumer i at period t first chooses which camera to buy, $B_{it} = 1,...,J_t$ (symbol B for "*buy*"), or not buying any new camera (and keeping the existing camera), in which case $B_{it} = 0$. The consumer can purchase any camera that has been announced, hence the time subscript in J_t. Buying a camera updates the existing camera in use, which we denote $C_{it} = 1,...,J_t$, and follows

$$C_{it} = \begin{cases} C_{it-1} & \text{if } B_{it} = 0 \\ B_{it} & \text{if } B_{it} > 0. \end{cases}$$

To consume a picture the consumer must produce it. In each period, the consumer makes a binary decision on whether to take a picture or not, $D_{it} = 0,1$. I do not model the number of pictures.[20] Apart from generating consumption utility, picture taking also generates a noisy signal – from which learning by doing takes place *a la* Jovanovic and Nyarko (1996) (hence symbol D for "*do*").

The timing of the decision problem is as follows. In each period t, the consumer first decides which camera to buy, B_{it}. If a different camera is purchased, switching cost is immediately effective on her human capital. Conditional on the latest camera at hand, the consumer then decides whether to take a picture with this camera. When a picture is taken, the consumer observes a signal that is informative on the ideal way of taking pictures – in other words, a picture serves as both a consumption good *and* a noisy signal for learning by doing.

[20] If I model the decision of the number of pictures to take in each period, I will run into two risks. First, there are a lot of variations in picture taking activities, for example whether one is on holiday or on a trip. Taking these into account will greatly complicate the estimation procedure, yet providing little insight to the central problem. Secondly, data-wise, modeling the number of pictures one takes is at risk of selection: it assumes that the observed pictures one uploads are all that he took or are selected using a fixed criteria.

4.3 Utility

With an abuse of notation, let us say that consumer i *previously* owned camera j^0 but now processes camera j. We allow for $j = j^0$ if she does not purchase *or* purchases the same camera. In the current period, she derives utility from camera specific characteristics for owning and purchasing a camera, X_j, as well as the quality of pictures she produces, $D_{it}Q_{ijt}$ – that is, consumption of pictures is made only when a picture is produced. Her *ex post* flow utility is specified as:

$$u_{ibdt}^{post} = \sum_j \left\{\theta \quad \log(P_{jt}) \quad \mathbf{1}(B_{it} = j) + X_j\beta \quad \mathbf{1}(C_{it} = j)\right\} + D_{it}\left\{\alpha \quad Q_{ijt} + \kappa\right\} + \varepsilon_{ibdt} \tag{1}$$

where D_{it} is the choice of taking pictures, α captures the utility from consuming a unit of picture quality,[21] κ captures the cost incurred from producing a picture (apart from its consumption value and gain to learning by doing), θ captures the (dis)utility from expenditure when a new camera is purchased, and β captures the utility from owning the camera j – which is a vector of characteristics specific to j. Note that α, θ, κ and β can be specified as random coefficients; however, as will be discussed later, specifying a finite mixture of random coefficients does not improve model fit.

4.4 Production of Pictures and Human Capital

The consumer decides *how* he uses his camera to produce pictures. If she is capable of using a given camera in a more "correct" way – so as to produce higher quality pictures, she is then considered to have a higher level of human capital. To formalize this idea, and to be able to capture the evolution of observed picture quality, I follow the Bayesian learning by doing framework by Jovanovic and Nyarko (1996) (applied in Foster and Rosenzweig (1995) and others), which will eventually derive a reduced form that captures the evolution of picture quality.

For a consumer i, who has decided to produce a picture at time t, there is an *ideal "method", Y_{ijt}*, from which, if implemented exactly, a picture of the highest possible quality (given the camera) will be produced. However, the ideal method is *ex ante* unknown to the consumer – in particular, each photographic scenario ($i \quad t$ combination) potentially requires a different method, although they are correlated within

[21]For simplicity, utility is assumed to be linear in quality, as in Jovanovic and Nyarko (1996). This provides significant simplification in the analytical derivation later, and I do not see clear benefit of going beyond this assumption. I also assume away resale value of any camera k.

the equipment j. Hence, not knowing the exact Y_{ijt}, the individual can only choose a method Z_{it} such that, according to her belief about Y_{ijt}, Z_{it} maximizes the expected picture quality produced.[22]

To fix ideas, think of picture quality being determined by – besides the camera quality – whether the camera is in focus or not. Let Y_{ijt} represent the ideal point of focus, i.e., the exact distance that the camera *should* focus on, and Z_{it} the user's chosen point of focus. One can then imagine that the picture quality would be lower than ideal, if it is out of focus ($Z_{it} \neq Y_{ijt}$), regardless of whether the camera is front focused ($Z_{it} < Y_{ijt}$) or back focused ($Z_{it} > Y_{ijt}$). For simplicity of the solution, let the picture quality be a quadratic loss function:

$$Q_{ijt} = \underbrace{\gamma_j}_{\substack{\text{technology} \\ \text{for camera } j}} \underbrace{\left[1 - (Y_{ijt} - Z_{it})^2\right]}_{\text{penalty for out of focus}} \tag{2}$$

where $\gamma_j > 0$ is a scalar that denotes the *objective* physical quality of the camera – which is fixed to camera j.

The ideal focusing distance Y_{ijt} is not known with certainty: let Ψ_{it} denote all prior information the consumer processes *at the beginning of* period t. Conditional on all available information, the consumer decides the actual focusing distance Z_{it}, by maximizing the expected picture quality in this period:[23]

$$\max_{Z_{it}} E\left[\gamma_j \left(1 - (Y_{ijt} - Z_{it})^2\right) | \Psi_{it}\right].$$

The solution to minimizing a quadratic loss is to choose the conditional expectation of the target, $Z_{it} = E\left[Y_{ijt} | \Psi_{it}\right]$. Therefore, one can write the expected picture quality, taking into account the agent's optimizing decision in picture taking, as

$$\begin{aligned} E\left[Q_{ijt} | \Psi_{it}\right] &= \gamma_j \left(1 - E\left[(Y_{ijt} - E\left[Y_{ijt} | \Psi_{it}\right])^2 | \Psi_{it}\right]\right) \\ &= \gamma_j \left(1 - Var\left(Y_{ijt} | \Psi_{it}\right)\right) \end{aligned} \tag{3}$$

[22] It is crucial that choice of z_{it} does not have dynamic implications. This will turn out to be true in the next section.

[23] When utility is linear in picture quality, and that Z_{it} does not have dynamic implication, this is equivalent to maximizing consumer utility. I do not see the benefit of relaxing these two assumptions because I only aim at deriving an analytical reduced form model of picture quality.

and I denote one minus the posterior variance, as the consumer human capital.

4.5 Learning by Doing and Switching Cost

To characterize how human capital evolves over time and degrades at camera switching, I will impose a structure on the target Y_{ijt}. For concreteness, think of Y_{ijt} as in the previous example, the ideal point of focus. For each scenario one faces, the ideal point of focus might be different; yet, there might be a systematic component in the way of focusing a given (type of) camera. Specifically, let the target Y_{ijt} be decomposed by a mean target and an independent and identically distributed (IID) shock

$$Y_{ijt} = \overline{Y}_{ij} + \omega_{ijt}$$

where $\overline{Y}_{ij}$ is a individual-camera specific fixed component that only depends on the type of camera one uses at t, and ω_{ijt} is a individual-time specific shock with $E[\omega_{ijt}|\overline{Y}_{ij}] = 0$.

In addition, θ_{ij} is fixed only given a camera j, but changes when the individual switches to another type of camera. For example, when switching to an SLR from a compact camera, the focusing system changes completely.[24] Formally, if the individual switches from j' in $t-1$, to j in t, then

$$\overline{Y}_{ij} = \overline{Y}_{ij'} + \eta_{ijj't}, \tag{4}$$

where the random variable $\eta_{ijj't}$ captures the additional uncertainty one faces, when given a camera with a unfamiliar focusing system. Substitute (4) into (3), we have:

$$\begin{aligned} Var(Y_{ijt}|\Psi_{it}) &= Var(\overline{Y}_{ij}|\Psi_{it}) + \sigma^2_\omega \\ &= \underbrace{Var(\overline{Y}_{ij'}|\Psi_{it}) + \sigma^2_\omega}_{\text{prior variance before switching}} + \underbrace{\sigma^2_{jj'}}_{\text{switching cost}} \end{aligned} \tag{5}$$

where the last term – variance of $\eta_{ijj't}$ – characterizes the switching cost. Note that I explicitly allow for the switching cost term to depend on the product pairs.

Next, after taking a picture, the consumer observes the ideal target the Y_{ijt}, from which she will reflect

[24] A compact camera uses a "point and shoot" style where one half-clicks the shutter release button to focus, but is generally slow; while an SLR requires one to select one among the few predetermined focusing areas when focusing, and usually supports manual focus.

what has been done wrong and can be improved in the next photographic scenario. In other words, Y_{ijt} serves as a signal for the consumer to Bayesian update her prior variance of the mean ideal target, θ_{jt}:

$$\frac{1}{\text{Var}\left(\overline{Y}_{ij}\, j\, \Psi_{it+1}\right)} = \frac{1}{\text{Var}\left(\overline{Y}_{ij}\, j\, \Psi_{t}\right)} + \frac{D_{it}}{\sigma_{\omega}^{2}}. \tag{6}$$

To ease notation, denote the *prior* variance $V_{it} = Var\left(\overline{Y}_{ij^0}\, j\, \Psi_{it}\right)$ as the state variable that captures individual human capital given the previously owned camera j^0, at the beginning of period t. We can now rewrite the (reduced form) production function and state transition as:

$$E\left[Q_{ijt} j C_{it}, C_{it\quad 1}, V_{it}\right] = \gamma_j \quad \left(1 \quad \left(V_{it} + \sigma_{jj^0}^{2} \quad 1\left(C_{it} = j; C_{it\quad 1} = j^0\right) + \sigma_{\omega}^{2}\right)\right) \tag{7}$$

and

$$V_{it}^{\quad 1} = \left(V_{it\quad 1} + \sigma_{j^0 j^{00}}^{2} \quad 1\left(C_{it\quad 1} = j^0; C_{it\quad 2} = j^{00}\right)\right)^{\quad 1} + \frac{D_{it\quad 1}}{\sigma_{\omega}^{2}}. \tag{8}$$

For notational simplicity, denote $\tilde{V}_{it} = V_{it} + \sigma_{jj^0}^{2} \quad 1\left(C_{it} = j; C_{it\quad 1} = j^0\right) + \sigma_{\omega}^{2}$ as the prior variance of target Y_{ijt}. I will use V_{it} and $\tilde{V}_{it}$ to represent the term "human capital" interchangeably, since they are one-to-one. The second equation captures the transition of prior to posterior variance, when both switching and learning takes place in period $t \quad 1$.

4.6 Dynamic Discrete Choice

In each period t, given the level of human capital and the camera she previously uses, the consumer first decides which camera to purchase (or not purchasing anything), and then whether to take a picture. In case when a picture is produced, she consumes the picture quality and updates her prior human capital, which concludes the period.

Before realization of the picture quality, the consumer have the *ex ante* flow utility, defined on the expected picture quality:

$$
\begin{aligned}
u_{ibdt} &= \sum_j \left\{ \theta \ \log(P_{jt}) \ \mathbf{1}(B_{it}=j) + X_j\beta \ \mathbf{1}(C_{it}=j) \right\} + D_{it}\left\{ \alpha_i \mathbb{E}\left[Q_{ijt}|C_{it}, C_{it-1}, V_{it}\right] + \kappa_i \right\} + \varepsilon_{ibdt} \\
&= \sum_j \left\{ \theta \ \log(P_{jt}) \ \mathbf{1}(B_{it}=j) + X_j\beta \ \mathbf{1}(C_{it}=j) \right\} + D_{it}\left\{ \alpha_i \ \gamma_j \left(1 - \tilde{V}_{it}\right) + \kappa_i \right\} + \varepsilon_{ibdt} \qquad (9) \\
&= u^I\left(B_{it}, C_{it-1}, \{X_j\}_{j\in J_t}, \{P_{jt}\}_{j\in J_t}\right) + u^{II}\left(B_{it}, D_{it}, C_{it-1}, V_{it}\right) + \varepsilon_{ibdt}
\end{aligned}
$$

In the last line, the first term denotes utility from camera purchase and ownership (in short, u^I_{ib}), and the second term the utility from picture production (u^{II}_{ibd}).

Having rational expectations over the transition process of state variables (i.e. camera ownership and human capital), and future decisions, the consumer makes choices B_{it} and D_{it} by maximizing the expected discount sum of future utility

$$
\max_{B_{it}, D_{it}} \mathbb{E}\left[\sum_{\tau=t}^{\infty} u_{ibd\tau}|V_{it}, C_{it-1}, \mathbf{X_t}, \mathbf{P_t}\right] \qquad (10)
$$

conditional on the beginning-of-period human capital V_{it}, camera C_{it-1}, price of all cameras $\mathbf{P_t} = \left\{P_{jt}\right\}_{j\in J_t}$ and other characteristics $\mathbf{X_t} = \left\{X_j\right\}_{j\in J_t}$. To simplify notation, denote the state space

$$
\mathbf{S_t} = \left(V_{it}, C_{it-1}, \mathbf{X_t}, \mathbf{P_t}\right).
$$

One can write probability of observing a time series of decisions for a given individual, as the joint distribution of

$$
\begin{aligned}
&\Pr\left(B_{i1}=b_1, D_{i1}=d_1; ...; B_{iT}=b_T, D_{iT}=d_T|V_{i1}; C_{i0}; \mathbf{S_1}, ..., \mathbf{S_T}\right) \\
&= \prod_{t=1}^{T} \Pr\left(B_{it}=b_t, D_{it}=d_t|V_{it}, C_{it-1}, \mathbf{S_t}\right) dF\left(V_{i1}\right) \qquad (11)
\end{aligned}
$$

To tractably characterize the above decision process, I make four assumptions that makes numerical characterization of the choice probabilities feasible. The first two are standard in the dynamic discrete choice literature since Rust (1987). I assume the infinite horizon decision problem is stationary, i.e.:

Assumption 1 For any two periods t and t',

$$
\Pr\left(B_{it}=b, D_{it}=d|\mathbf{S_t}\right) = \Pr\left(B_{it'}=b, D_{it'}=d|V_{it'}, C_{it'-1}, \mathbf{S_{t'}}\right).
$$

And I impose a distribution assumption on the utility shock ε_{ibdt}:

Assumption 2 The utility shock ε_{ibdt} can be decomposed into

$$\varepsilon_{ibdt} = (1-\lambda)\,\varepsilon^{I}_{ibt} + \lambda\,\varepsilon^{II}_{idt}$$

where ε^{I}_{ibt} and ε^{II}_{idt} follow IID type 1 extreme value distribution.

Denote the expected maximum of current and future utility from making choice combinations $B_{it} = b$ and $D_{it} = d$ as U_{ibd}, and thus the expected maximum of current and future utility condition on choice B_{it} but before D_{it} is made is $\mathbb{E}\left[\max_{D_{it}} U_i\left(B_{it}, D_{it}, \mathbf{S_t}\right) | B_{it}, \mathbf{S_t}\right]$, and we have:

$$\begin{aligned}
\Pr\left(B_{it} = b, D_{it} = d | \mathbf{S_t}\right) &= \int \Pr\left(D_{it} = d | B_{it}, \mathbf{S_t}\right)\ \Pr\left(B_{it} = b | \mathbf{S_t}\right) dF\left(i\right) \\
&= \int \frac{\exp\left(U_{ibd}/\lambda\right)}{\sum_{d=0,1}\exp\left(U_{ibd}/\lambda\right)}\ \frac{\exp\left(\mathbb{E}\left[\max_{D_{it}} U_{ibd} | B_{it}, \mathbf{S_t}\right]/(1-\lambda)\right)}{\sum_{b=0,\dots,J_t}\exp\left(\mathbb{E}\left[\max_{D_{it}} U_{ibd} | B_{it}, \mathbf{S_t}\right]/(1-\lambda)\right)} dF\left(i\right).
\end{aligned}$$

Where the expected maximum current and future utility from making b and d (or the *choice specific value function*) is defined recursively by the Bellman equation:

$$\begin{aligned}
U_{ibd} &\equiv U_i\left(B_{it}, D_{it}, \mathbf{S_t}\right) \\
&= u^{I}_i\left(B_{it}, C_{it-1}, \mathbf{X_t}, \mathbf{P_t}\right) + u^{II}_i\left(B_{it}, D_{it}, C_{it-1}, V_{it}\right) + (1-\lambda)\,\varepsilon^{I}_{ibt} + \lambda\,\varepsilon^{II}_{idt} \\
&\quad + \delta\ \mathbb{E}\left[\max_{B_{it+1}, D_{it+1}} U_i\left(B_{it+1}, D_{it+1}, \mathbf{S_{t+1}}\right) | B_{it}, D_{it}, \mathbf{S_t}\right]. \qquad (12)
\end{aligned}$$

4.7 Curse of Dimensionality

Solving for the choice probabilities hinges on the solution of the choice specific value function $U_i\left(B_{it}, D_{it}, \mathbf{S_t}\right)$, which cannot be obtained analytically. One commonly used method in the dynamic discrete choice literature(Rust, 1987, 1994), is to numerically solve U_i as a nonparametric function on a discrete state space, by iterating the Bellman Equation (12). This brings up the curse of dimensionality problem, arose from a large product space. On the one hand, there is a large set of production function parameters to be estimated in the first place. We need to estimate J camera technology parameters γ_j, and $J \times (J-1)$ pairwise switching costs $\sigma^2_{jj'}$. This is impossible given that there are over 500 products in the data-set. On the other hand, the need to solve U_i over a state space, that is the size of the Cartesian product of current and past

product space, human capital (which is a continuous variable), price and characteristics of *all* products, and a binary decision of picture taking. Solving for U_i over such a state space is impossible.[25]

As a practical solution to restrict the parameter space, I impose that picture quality is systematically different between two types of cameras: a compact camera (denote $\overline{C}_{it} = 1$) and an SLR ($\overline{C}_{it} = 2$); but not systematically different within type. This is to impose that camera technology is summarized by two type-specific parameters, $\overline{\gamma}_k, k = 1,2$, while the pairwise switching cost summarized by $\sigma^2_{kk'}$ for an individual switching from type k' to k. One naturally restricts that within-type switching cost is zero, or $\sigma^2_{kk} = 0$.

While the above restrictions reduces the parameter space, it does not reduce the state space, and one cannot simply average price and product characteristics within category. Hence, and because of the dynamic nature, utility still depends on the identity of the past and current product. Without mechanically reducing the product space, or making naive assumptions on dynamics and rational expectations, I provide two alternative assumptions to further reduce the state space relevant to the dynamic decision problem, using a method similar to Hendel and Nevo (2006).

I first assume that the purchase-relevant utility shock, ε^I_{ibt}, can be further decomposed into two IID type 1 extreme value errors:

Assumption 3 For $b = 1,...,J_t$:

$$\varepsilon^I_{ibt} = \varepsilon^i_{ikt} + \varepsilon^{ii}_{ijt} j_k$$

where $k = 1,2$, and ε^i_{ikt} and $\varepsilon^{ii}_{ijt} j_k$ are IID type 1 extreme value.

Naturally, $\varepsilon^I_{i0t} = \varepsilon^i_{i0t}$. Denote the choice over camera categories $\overline{B}_{it} = 0,1,2$ where 0 stands for no purchase, and the static maximum utility from purchasing from category k as Γ_{ikt}. A natural restriction for the transition of Γ_{i0t} is that, if no camera is purchased in t, Γ_{i0t} should be equal to the utility from owning the previously purchased camera. In addition, we need to restrict the transition of Γ_{ikt} for $k \neq 0$ (instead of the traditional restriction on the transition of camera characteristics X and P):

Assumption 4 The expected maximum static utility from purchasing a camera in category $k = 1,2$, Γ_{ikt}, follows a first order Markov process.

With all 4 assumptions, it is possible to establish another Bellman equation where the camera-category specific value function is defined on the reduced product category space only. This is formalized in the

[25] Even without a large product space, the dimension of product characteristics $\mathbf{X_t}$ alone will make solving for U_i difficult.

following Lemma:

Lemma 1. *Under Assumptions 1-4, and when production-side parameters are restricted to $\overline{\gamma}_k$ and $\sigma_{kk'}$, the dynamic optimization problem (10) given the utility function (9), production function (7), transition process of human capital (8), and discount factor δ, has an associated Bellman equation:*

$$
\begin{aligned}
\overline{U}_i\left(\overline{B}_{it}, D_{it}, V_{it}, \overline{C}_{it-1}, \Gamma_{it}\right) &= \theta \cdot \log\left(\overline{P}_k\right) \cdot \mathbf{1}(k \neq 1) + \overline{u}^{II}_{ikd} + (1-\lambda)\,\varepsilon^{i}_{ikt} + \lambda\,\varepsilon^{II}_{idt} + \Gamma_{it} \\
&\quad \delta \cdot \mathbb{E}\left[\max_{\overline{B}_{it+1}, D_{it+1}} \overline{U}_i\left(\overline{B}_{it+1}, D_{it+1}, V_{it+1}, \overline{C}_{it}, \Gamma_{it}\right) | B_{it}, V_{it}, \overline{C}_{it-1}, \Gamma_{it}\right], \quad (13)
\end{aligned}
$$

where

$$
\begin{aligned}
\Gamma_{ik}\left(C_{it-1}, \mathbf{X_t}, \mathbf{P_t}\right) &= \mathbb{E}\Bigg[\max_{B_{it}|k} \sum_{j \in k} \\
&\quad \left\{\theta \cdot \log\left(\frac{P_{jt}}{\overline{P}_k}\right) \cdot \mathbf{1}(B_{it}=j) + X_j\beta \cdot \mathbf{1}(C_{it}=j) + (1-\lambda)\,\varepsilon^{ii}_{ijt}|_k\right\} | k, C_{it-1}, \mathbf{X_t}, \mathbf{P_t}\Bigg] \quad (14)
\end{aligned}
$$

And

$$
\Gamma_{it} \equiv \{\Gamma_{ikt}\}_{k=0,1,2}.
$$

The proof will be detailed in the Appendix.

4.8 2 Step Estimator

4.8.1 The first step: static camera choice

Clearly, to obtain a solution of the dynamic decision problem in the reduced state space, (13), we need to first compute the state variable Γ_{it}. This is done by a first step estimator on the static decision problem

$$
\max_{B_{it}|k} \sum_{j \in k} \left\{\theta \cdot \log\left(\frac{P_{jt}}{\overline{P}_k}\right) \cdot \mathbf{1}(B_{it}=j) + X_j\beta \cdot \mathbf{1}(C_{it}=j) + (1-\lambda)\,\varepsilon^{ii}_{ijt}|_k\right\}
$$

i.e. conditional on choosing a camera-category decision $k = 1, 2$, which camera model does the consumer purchase. Note that because we need to estimate this static decision problem separately, we cannot allow for individual random coefficients to be correlated with random coefficients in the second step. Hence, I only estimate the static decision problem for a representative individual. We could, however, separately

estimate the static model to allow for product-category specific coefficients. I also re-normalize the scale of error term to allow $(1 \quad \lambda)\, \varepsilon^{ii}_{ijt} j_k$ to be extreme value type 1.

Estimating the static decision parameters separately will give the expected maximum utility for purchasing in each category, Γ_{ikt} for $k = 1,2$. I then compute the *utility* from owning the previous camera, Γ_{i0t}, as linear projection of the static discrete choice model:

$$\Gamma_{i0}\left(C_{it \quad 1} = j, X_{jt}\right) = X_j \hat{\beta}^k;$$

and re-normalize the inclusive values of purchasing so that the static utility of *not* purchasing is zero, or:

$$\tilde{\Gamma}_{ikt} = \Gamma_{ikt} \quad \Gamma_{i0t}.$$

Without loss of generality, this re-normalization reduces the dimension of state space by 1. Finally, by assuming Γ_{ikt} is first order Markov, I estimate a reduced form transition process where

$$\tilde{\Gamma}_{ikt} = b^k_0 + b^k_1 \tilde{\Gamma}_{i1t \quad 1} + b^k_2 \tilde{\Gamma}_{i2t \quad 1} + \zeta_{ikt}. \tag{15}$$

And the individual will take the linear prediction $\hat{\Gamma}_{ikt+1}\left(\tilde{\Gamma}_{i1t}, \tilde{\Gamma}_{i2t}\right)$ to form their expectation of future inclusive values.

4.8.2 The second step

Given estimates of the static discrete choice problem and the Markov transition processes, I can now jointly estimate the production function, the human capital evolution process, and the utility parameters related to the dynamic decisions. Alternatively, one could first estimate the production-side parameters, and then maximize the partial likelihood on the dynamic discrete choice problem.

To be able to characterize the distribution of *observed* picture quality, I impose that observe quality is the sum of expected picture quality and a measurement error:

$$Q_{ijt} = \mathbb{E}\left[Q_{ijt} j \overline{C}_{it}, \overline{C}_{it \quad 1}, V_{it}\right] + \xi_{it},$$

where learning by doing and production function is characterized by a dimensionality-reduced version of

Equations (7) and (8), namely:

$$E\left[Q_{ijt} | \overline{C}_{it}, \overline{C}_{it-1}, V_{it}\right] = \overline{\gamma}_k \left(1 - \left(V_{it} + \sigma^2_{kk'} \, 1\left(\overline{C}_{it} = k; \overline{C}_{it-1} = k'\right) + \sigma^2_{\omega}\right)\right) \tag{16}$$

and

$$V_{it}^{-1} = \left(V_{it-1} + \sigma^2_{k'k''} \, 1\left(\overline{C}_{it-1} = k'; \overline{C}_{it-2} = k''\right)\right)^{-1} + \frac{D_{it-1}}{\sigma^2_{\omega}} \tag{17}$$

and the measurement error ξ_{it} follows a logistic distribution with scale ψ_q and location 0.

Given estimates from the production function, one estimates the parameters in the utility function (9) that are related to the decisions of purchasing any camera in a category and taking a picture. For a set of candidate parameter, the Bellman equation (13) can be iterated to solve for a nonparametric choice specific value function $\overline{U}_{ikd}$, which can be used to construct a likelihood function on the choice probabilities of picture taking and camera purchases. By construction of the data, purchasing a camera implies taking a picture in the given period.[26] Hence, I eliminate the possibility of camera purchases with no picture taken, and leads to the following partial likelihood function:

$$\begin{aligned} \log\left(L^{II}\right) &= \sum_{i,t} \sum_{k=1,2} \left\{ \log\left(\frac{\exp\left(\mathbb{E}_t\left[\max_{D_{it}} \overline{U}_{ikd}\right]/(1-\lambda)\right)}{\sum_{k=1,2} \exp\left(\mathbb{E}_t\left[\max_{D_{it}} \overline{U}_{ikd}\right]/(1-\lambda)\right)} \right) \mathbf{1}\left(\overline{B}_{it} = k\right) \right\} + \\ & \sum_{i,t} \sum_{d=0,1} \left\{ \log\left(\frac{\exp\left(\overline{U}_{i0d}/\lambda\right)}{\sum_{d'=0,1} \exp\left(\overline{U}_{i0d'}/\lambda\right)} \right) \mathbf{1}\left(D_{it} = d; \overline{B}_{it} = 0\right) \right\}. \end{aligned}$$

Alternatively, one could jointly estimate the parameters by maximizing the likelihood on the distribution of $\overline{B}_{it}$, D_{it} and Q_{it}.

4.8.3 Initial conditions and the discount factor

This paper will remain silent on how the consumers choose their first camera in the initial period. While it might be of interest to analyze the initial decision of an individual who rationally expects future switching cost, this is not possible in my data and framework.

For initial conditions on the human capital, I take the projected fixed effects from the regression on

[26] Which also seems realistic.

the number of views as in Equation (21), denoted μ_i, and then define the initial condition to be

$$V_{i1} = V_{max} \quad \frac{\mu_i \quad \text{min}_i(\mu_i)}{\text{max}_i(\mu_i) \quad \text{min}_i(\mu_i)}. \tag{18}$$

In the empirical implementation, I take $V_{max} = 0.8$ to limit the size of state space.[27] The rationale behind this parametrization is that those who have higher picture quality *throughout* the course of my observation, are in possession of better knowledge to photography at the very beginning. As an alternative interpretation, $\mu_i^0 s$ could reflect differences individual-specific popularity rather than initial human capital endowments. Nevertheless, since I can trace every user to the beginning of their usage of Flickr (but I cannot trace back to the start of their hobby in photography), it is much more plausible to assume that popularity is identical at start-up, and such initial differences reflect heterogeneity in human capital endowments.

Conventional in empirical works using the dynamic discrete choice framework, the discount factor, δ, is not identified. I use a monthly discount factor of $\delta = 0.95$. While this is much smaller than one would infer from the market interest rate, for a camera market, it predicts that people mentally depreciates 85% of the value of a camera in 3 years.

5 Estimation Results

5.1 Static Utility Parameters

This section discusses estimation results and model fit. In the first step, I estimate a static discrete choice problem conditional on the purchase decisions of camera categories. Specifically, individual's utility of purchasing and owning a camera depends on the price of the camera, quadratic specifications of the number of periods since the camera has been announced and camera resolution (megapixels), manufacturer's announced price (as a proxy to unobserved quality), and so forth. For compact cameras, I found that the age of camera confounds with prices; hence, I do not use this covariate for the this type of cameras.

Table 5 presents the results. Except for the coefficient on *dpreview.com* rating (β_8), the signs of parameter estimates are intuitive. Specifically, the consumers has a tendency to select camera models that are higher in megapixels and lower in prices. For choices of SLR cameras, the consumer likes more recent

[27]Choosing V_{max} between 0.8 and 1.2 does not bring discernible difference to the results.

Table 5: Estimates: Static Utility Parameters

	Compact	std. err.	SLR	std. err.
$\log(price)$ (β_1)	-1.0621	0.1152	-0.4970	0.5288
Age of the camera (β_2)	NaN	NaN	-0.7716	0.2236
Age squared (β_3)	NaN	NaN	0.0763	0.0256
Megapixel (β_4)	0.6856	0.0356	0.2142	0.0349
Megapixel squared (β_5)	-0.0432	0.0019	-0.0060	0.0023
$\log(MSRP)$ (β_6)	-0.0158	0.0688	0.0012	0.0403
Larger sensor (β_7)	0.5723	0.1343	0.0186	0.3138
Rating: good and very good (β_8)	-1.2470	0.1336	-0.0600	0.2693
Brand: Canon (β_9)	1.4165	0.0582	1.1291	0.8977
– Nikon (β_{10})	-0.9440	0.1298	1.2282	1.0035
– Sony (β_{11})	1.2384	0.1516	-1.0581	0.2196
– Pentax (β_{12})	-1.7757	0.0514	-0.7700	1.3748
– Other (β_{13})	0.2361	0.0431	-0.5296	1.0327

Note: Sub-sampled standard errors (from 50 iterations) are provided.

Table 6: Estimates: Markov Evolution of the Inclusive Values

	ω_{i1t}	std. err.	ω_{i2t}	std. err.
Constant (θ_c)	0.3513	0.0148	-0.2238	0.0151
ω_{i1t-1} (θ_1)	0.9405	0.0047	0.1270	0.0048
ω_{i2t-1} (θ_2)	0.0154	0.0048	0.8261	0.0048
R-squared	0.9255	0.0000	0.9215	0.0000

Note: Asymptotic standard errors are provided.

cameras, while having an aversion to higher prices. This suggests a tradeoff between waiting for prices to decrease, and having the latest camera.

For the magnitude of price coefficients, one finds that the estimates as presented in Table 5 is rather similar to the literature on discrete choice of durable goods, e.g. Song and Chintagunta (2003) – despite lacking dynamics and heterogeneity. To more closely look at the magnitude of price coefficients, I compute the implied elasticities as in a standard Logit model, for all observations which we used in the first step estimation.[28] That is, the elasticity computed here is conditional on the individual deciding to purchase one of the cameras in the given category. Then, I average the elasticities given the identity of camera, as the *average* elasticity for a given camera model, shown in Figure 3. For example, for the median compact camera, a one percent *decrease* in the price will trigger a 5 percent *increase* in its purchase

[28] I.e. observations when an individual decides to switch cameras.

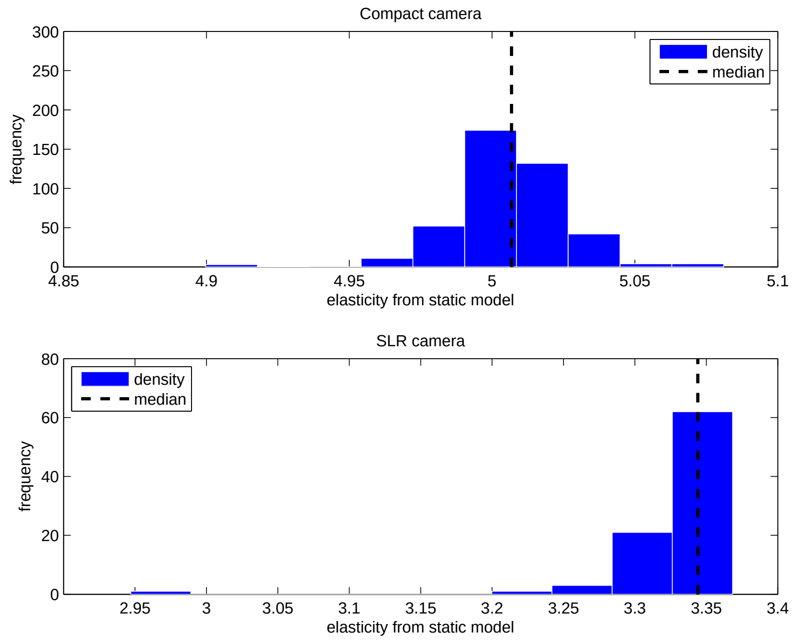

Figure 3: Predicted Static Price Elasticity

Note: Shows the distribution of (negative of) the average elasticities for a given camera model, conditional on the individual deciding to purchase one of the cameras in the category. Note that these are not "true" elasticities from the full model.

probabilities.

The brand dummies has intuitive size and magnitude as well. One finds that for SLRs, the high Canon and Nikon brand dummies explains their considerably higher market shares; while for compact cameras, Canon has a clear advantage over all other brands. In another version (not reported), I also added product dummies for the top 50 popular products. Despite increasing the computation burden, this does not bring discernible changes to the estimates the main parameters of interest.

I then estimate a reduced form first order Markov process of the projected inclusive values in Equation (15). As indicated by the R-squared, a first order Markov specification captures most of the variations in the inclusive values.

Table 7: Estimates: Production Function of Picture Quality

	separate est	std err	joint est	std err
Physical quality: Compact ($\overline{\gamma}_1$)	1.5347	0.1324	2.5292	0.1507
Physical quality: SLR ($\overline{\gamma}_2$)	2.3656	0.1434	2.8325	0.1684
Noise of signal – compact camera ($\sigma_{\omega,1}$)	0.9407	0.0093	0.9541	0.0036
– SLR ($\sigma_{\omega,2}$)	0.8359	0.0108	0.8691	0.0063
Switching cost to SLR ($\sigma_{1,2}$)	0.2653	0.0428	0.1102	0.0390
Switching cost to compact ($\sigma_{2,1}$)	0.0022	0.0017	0.0012	0.0008
Scale of error (σ_{ξ})	0.7453	0.0084	0.7450	0.0065

Note: Sub-sampled standard error (from 50 iterations) are provided.

5.2 Dynamic Parameters

5.2.1 Production Function and Learning by Doing

One can then estimate all parameters relevant to the dynamic discrete choice problem, which include parameters in the production function, evolution of human capital, and utility parameter relevant to picture quality or picture-taking. We can either estimate the production function and human capital evolution separately, or jointly with the dynamic discrete choice problem. In either case, I estimate the production function and evolution of human capital in Equations (16) and (17) using the data in periods $2,\ldots,T$ ($T = 120$ in units of months). Data in the first period is used to determine initial human capital.

First of all, there is significant vertical differentiation between the physical quality of the two types of cameras ($\gamma_2 > \gamma_1 > 0$): at a given level of human capital, an SLR is more productive than a compact camera in the quality of pictures it produces. Such difference is visually apparent when plotting the quality of pictures a consumer produces against his experience, as shown in Figure 4.

On the other hand, for a consumer using the camera, there is strong evidence on learning by doing and switching cost. According to the Bayesian learning structure and the estimates on $\sigma^2_{\omega,k}$, the model suggests that learning speed is initially very fast, but decreases dramatically as one progresses. Potentially, the learning speed is different between compact cameras and SLRs as well. In addition to that, there is considerable switching cost when the individual changes her compact camera into an SLR – but not otherwise. This suggests that human capital is backward compatible: an SLR is potentially more difficult to use, and hence requires considerably better knowledge in photography.

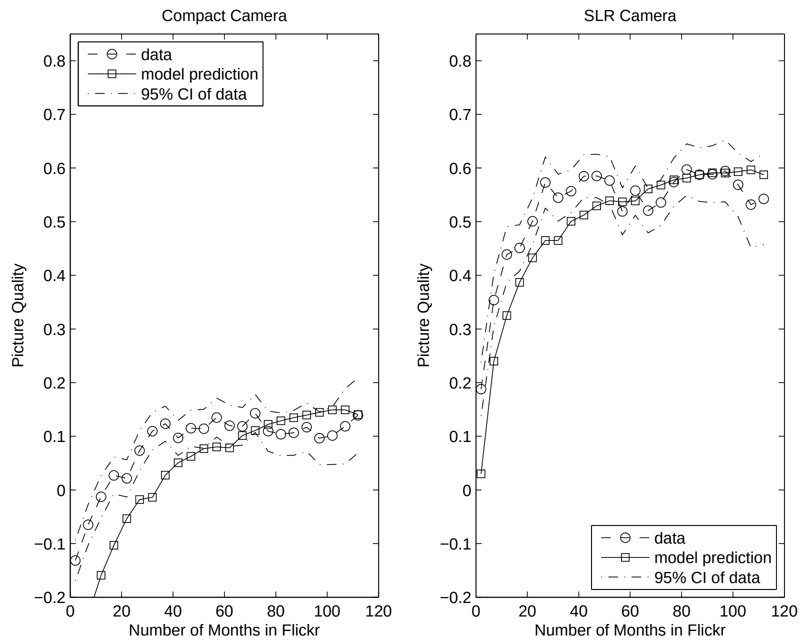

Figure 4: Model Fit on Picture Quality

Note: Produced by parameter estimates from jointly estimating the production and utility parameters, i.e. column 3 in Table 7.

I also examine model fit in 4. One finds that the simple production function and human capital evolution process can capture the basic shape of the picture quality process over time, with most of the model predictions lying in the 95% confidence interval of the data. Note that if we only fit the marginal distribution of picture quality Q_{it}, then it is not surprising to find the fit for picture quality to be much better.

Another way of looking at learning by doing and switching cost is through Figure 5. In this figure, the distribution of model-predicted human capital is shown at a few cross-sections. [29] The figure shows that human capital increases drastically in the first few periods, and virtually stops growing after the first *5 years*. The median of the human capital distribution thereafter is also the median human capital (0.95), which, as we will see in subsequent sections, has dramatically different choice patterns than the initial consumers (who are at lower human capital levels).

[29] To be precise, I plot 1 $Var(\theta_{ij}j\Psi_{it})$ instead of 1 $Var(\theta_{ij}j\Psi_{it})$ $\sigma^2_{\omega,k}$, because this allows me to keep the horizontal axis in the same range across products

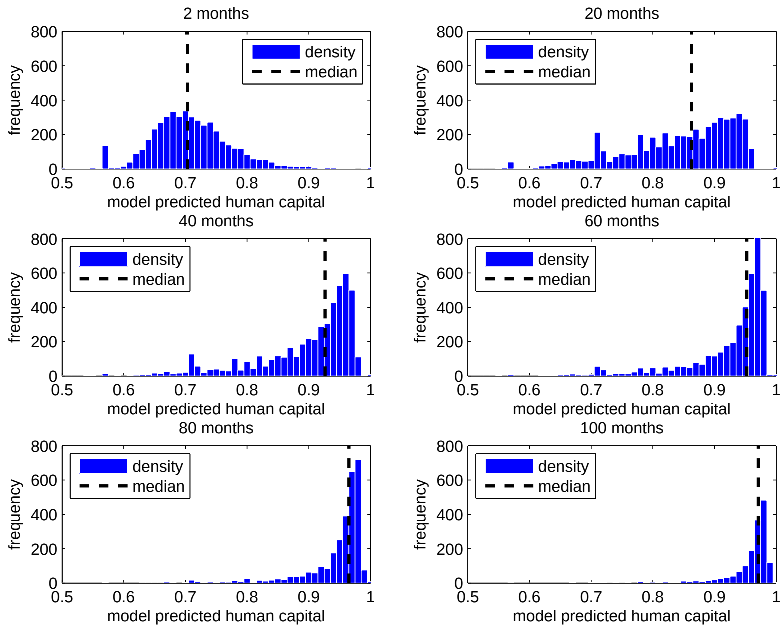

Figure 5: Distribution of Model-Predicted Human Capital

5.2.2 Dynamic Utility Parameters

I now turn to estimation of the discrete choice problem, given the projected inclusive values from static decisions. The results are presented in Table 8. One starking findings from the parameter estimates, is that the consumers in our sample weighs picture quality more than price. For example, if purchasing a new camera will improve the picture quality to such an extent that it will attract 100% more viewers,[30] the individual is willing to pay for (much) more than 100% in price.

To examine the fit for the entire model we have estimated, I simulate a counterfactual population with exactly the same initial conditions as in the first 10 periods in the data, and let them make camera purchasing and picture-taking decisions – as well as human capital evolutions – following the model estimates. Figure 6 depicts the counterfactual and observed share of SLR cameras in the sample, i.e. the predicted diffusion path of SLR cameras. It seems that the model fits the data rather well except for the final periods, when most of the individuals show missing data. Hence, this figure shows acceptable model fit.

[30]The reason for using such a metric is because picture quality is proxied by a function of log number of views. Note that this is *different* from assuming that consumers ultimately care about popularity.

Table 8: Estimates: Dynamic Utility Parameters

	separate est	std err	joint est	std err
Picture quality (α)	1.1524	0.2320	0.8009	0.1688
Effort – compact (c_1)	-0.1752	0.0711	-0.1359	0.0396
– SLR (c_2)	-0.7297	0.1587	-0.4969	0.1157
log $(\overline{price})$	-0.5504	0.0347	-0.2736	0.0518
Intercepts – compact to compact (τ_{11})	-0.3862	0.0774	-1.6373	0.1605
– compact to SLR (τ_{12})	0.8381	0.1855	-1.3165	0.2881
– SLR to compact (τ_{21})	-0.7101	0.0692	-2.2108	0.1655
– SLR to SLR (τ_{22})	0.6189	0.1643	-1.3040	0.2689
Scale of static incl value	0.0513	0.0084	0.0454	0.0103
Weight of utility shock (λ)	0.1103	0.0196	0.1094	0.0208

Note: See notes in Table 7.

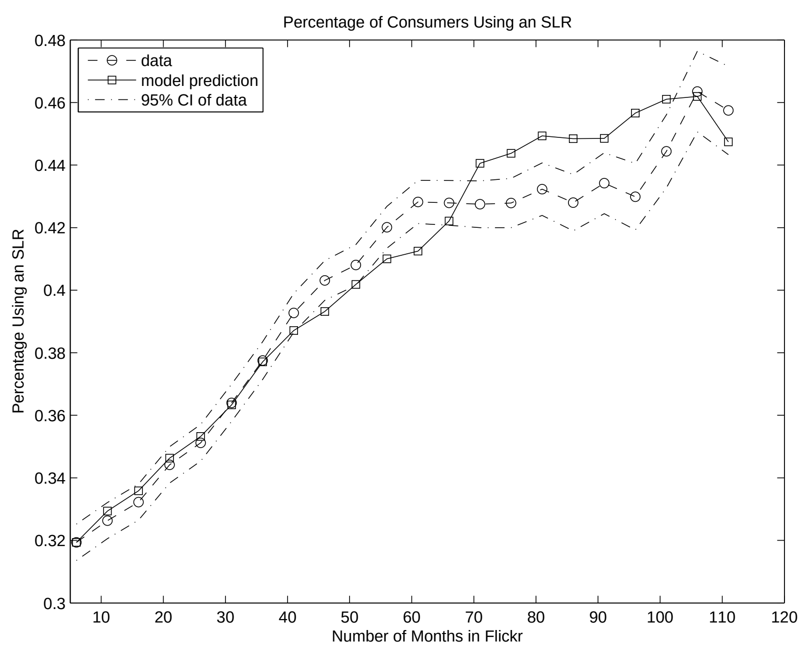

Figure 6: Model Fit on Camera Share

Note: The squares represents market share for SLR cameras, for a counterfactual population that shares the same initial conditions for the first 10 periods, and then makes their own decisions on picture taking and camera switching as predicted by the model. I use parameters produced by joint estimation (column 3 in Table 7 and 8), while using results in column 1 produces a very similar figure.

5.2.3 Random Coefficients

Up until now, the model and the estimates assumes that all unobserved heterogeneity in consumer decisions are orthogonal to the observables (i.e. they are captured in the $\varepsilon' s$). Despite the fact that the literature typically assumes that the model coefficients are random variables that have discrete distributions (Song and Chintagunta, 2003), and interpret different realizations of the same coefficient as "types" or "segments" of consumers, I am capturing individual heterogeneity by allowing for Markov heterogeneity in human capital and its path of evolution. It turns out, however, that there is no additional explanation power for my model if adding finite mixtures. It seems that all heterogeneity in my sample is captured by consumer human capital. Of course, this cannot be over-generalized since this paper only focuses on heavy users of Flickr, which are arguably the groups that values picture quality significantly more than average.

6 Policy Implications

We have demonstrated that when the consumer moves up the quality ladder of digital cameras, there is a nontrivial consumer switching cost, which might arise from the lack of user-friendliness in the more advanced SLR cameras. In this section, I explore a two-fold question: i) to what extent will the social planner be willing to invest in product-redesign, so as to improve the user-friendliness of an SLR camera; and ii) is there any supply side incentives to undertake such investments, in terms of the potential profit gains that a firm can enjoy? Essentially, the first question asks for the willingness to pay for such a hypothetical innovation for the consumers in my sample, while the second one examines their actual choice probabilities under the counterfactual policy.

6.1 Welfare for Eliminating Switching Cost

I first consider the welfare gain if the social planner decides undertake an innovation (assumed costless) to improve the user-friendliness of SLR cameras, so that there is no switching cost on human capital for all consumers. This welfare gain can be measured by the compensating variation for a policy that eliminates the switching cost – that is, we ask for the consumers' willingness to pay (WTP), in the counterfactual world with no switching cost, to set their discounted lifetime utility to the same level as in the real world.

I compute the compensating variation in terms of the price premium of SLRs, by first solving for

Table 9: Mean Counterfactual Outcomes from Reducing Switching Cost

	separate est	joint est
Equal utility	25.8214	14.0430
Equal choice probability	17.8455	10.1606

Note: The first number reads: as predicted by parameter estimates in column 1 in Tables 7 and 8, the mean WTP for elimination of the switching cost (i.e. the compensating variation) is equivalent to a 25.8% increase in the SLR prices.

inclusive values of buying SLR cameras in (14) when setting the new price, and then solving the Bellman equation (13) when setting the switching cost $\sigma^2_{kk^0} = 0$. One can find the compensating variation for each observation, by iterating the above procedure for different trial values of price premium.

The first row in Table (9) summarizes the mean WTP from all observations (individual-time) in terms of price premium of SLRs, under different specifications. The mean willingness to pay is equivalent to a 17% or 25% *permanent* increase in the SLR prices, or \$150 - \$220 dollar discounts when purchasing SLRs. In addition, there is considerable heterogeneity in the WTP distribution. As shown in figure 7, consumers' willingness to pay ranges from 0 to more than 50% of SLR prices. This largely depends on the consumer human capital, which is heterogeneous across and within individuals.

6.2 Will Firms Voluntarily Reduce Switching Cost?

Despite the heterogeneity in the willingness to pay across consumers, the previous section shows significant welfare gains from the improved user-friendliness of SLR cameras. On the other hand, would firms be willing to invest if product-redesign is costly? To investigate this question, I examine the SLR price premium a firm is able to charge in the counterfactual world, to maintain that the consumers' purchase probabilities of SLR cameras are equal to those in the real world.[31] This price premium gives a crude idea of how much more profit a firm would gain if it decides to develop products that have lower switching cost. The resulting distribution is plotted in Figure 8. As compared to Figure 7, a firm has lower incentive to undertake such an investment, compared to the social planner. This indicates that there are consumer surplus from reducing switching cost, that cannot be captured by the firm.

[31] In this experiment, I assume that there is a monopolist in the market. The results are very similar if I use the actual market structure.

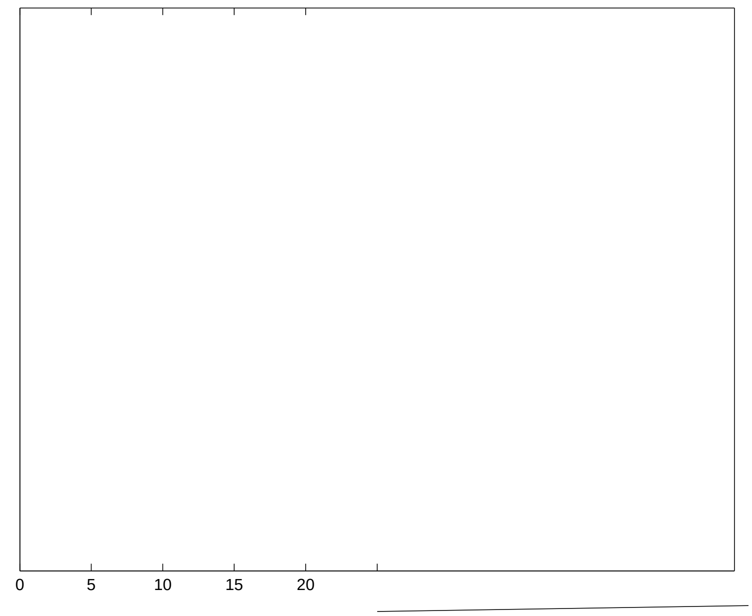

Figure 7: WTP for Reducing Switching Cost

Notes: Distribution of WTP across all individual-time. This figure is produced using the parameters from joint estimation of the model, i.e. column 3 in Tables 7 and 8.

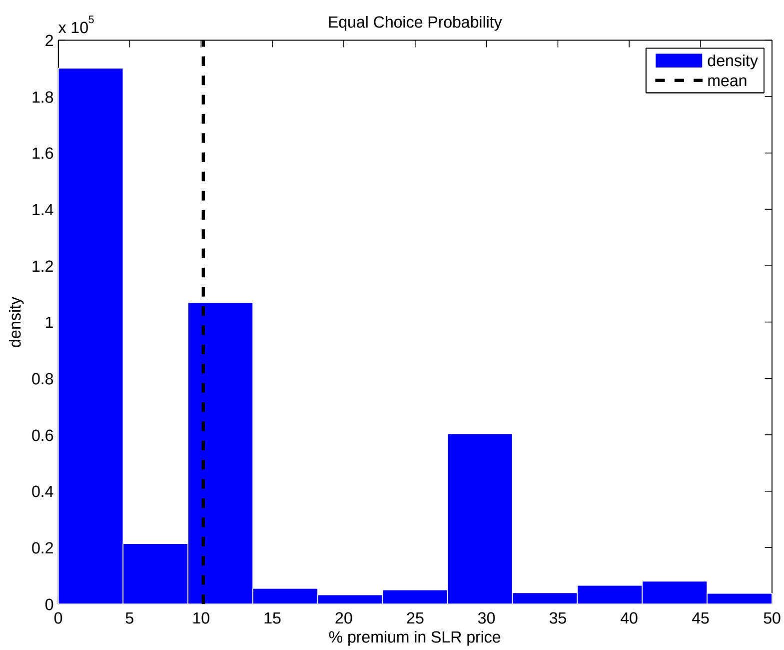

Figure 8: Price Premium for Reducing Switching Cost: Monopoly Case

Notes: Distribution of price premium that a monopolist can charge to equate consumers' choice probabilities towards SLRs. This figure is produced using the parameters from joint estimation of the model, i.e. column 3 in Tables 7 and 8.

7 Concluding Remarks

What internal factors can induce consumer switching, even when the environment is stable? While the majority of the marketing models explore the effect of familiarity (or formally, information) on product purchases or switching behavior, this paper formalizes the effect of expertise – or more precisely, consumers' product specific human capital. For consumers of digital cameras, I characterize human capital as the capability in producing high quality pictures, which proves to be a key state variable in explaining the forward-looking decisions of camera purchases and picture taking. Specifically, the empirical results suggest that consumer human capital can be improved as one gains experience by producing pictures (learning by doing), yet can also deteriorate as one switches to SLR cameras that are difficult to use, despite having higher potential in producing quality pictures.

The latter suggests a source of switching cost in human capital – arose from the lack of user-friendliness in the SLR product design. This is to say, if the SLRs remain equally capable of producing high quality pictures, but are improved in their user-friendliness to match that of a compact camera, then there is considerable space of welfare improvements. In the counterfactual experiments, consumers on average are willing to pay a 17%-25% price premium to the SLRs if such an innovation takes place. On the other hand, even though a firm who undertakes such an effort in product design cannot fully capture the consumer surplus, I show that a firm can charge 10%-14% price premium on each units of SLRs, to maintain the same quantity sold. This helps explain the recent emphasis on user-friendliness of the SLR cameras, and the new trend in production of hybrid cameras that share the ease of use among compact cameras.

References

Alba, Joseph W and J Wesley Hutchinson (1987), 'Dimensions of consumer expertise', *Journal of consumer research* pp. 411–454.

Arrow, K.J. (1962), 'The economic implications of learning by doing', *The review of economic studies* **29**(3), 155–173.

Becker, G.S. (1965), 'A theory of the allocation of time', *The economic journal* **75**(299), 493–517.

CEA (2012), 'The changing landscape of digital photography'.

Crawford, G.S. and M. Shum (2005), 'Uncertainty and learning in pharmaceutical demand', *Econometrica* **73**(4), 1137–1173.

Erdem, Tulin and Michael P. Keane (1996), 'Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets', *Marketing Science* **15**(1), 1–20.

Foster, A.D. and M.R. Rosenzweig (1995), 'Learning by doing and learning from others: Human capital and technical change in agriculture', *Journal of Political Economy* pp. 1176–1209.

Hendel, Igal and Aviv Nevo (2006), 'Measuring the implications of sales and consumer inventory behavior', *Econometrica* **74**(6), 1637–1673.

Hendricks, K. and A. Sorensen (2009), 'Information and the skewness of music sales', *Journal of Political Economy* **117**(2), 324–369.

Israel, Mark (2005), 'Services as experience goods: An empirical examination of consumer learning in automobile insurance', *The American Economic Review* **95**(5), 1444–1463.

Jovanovic, B. and Y. Nyarko (1996), 'Learning by doing and the choice of technology', *Econometrica: Journal of the Econometric Society* pp. 1299–1310.

Moshkin, N.V. and R. Shachar (2002), 'The asymmetric information model of state dependence', *Marketing Science* **21**(4), 435–454.

Osborne, M. (2007), 'Consumer learning, switching costs, and heterogeneity: A structural examination', *Quantitative Marketing and Economics* pp. 1–46.

Rust, J. (1987), 'Optimal replacement of gmc bus engines: An empirical model of harold zurcher', *Econometrica: Journal of the Econometric Society* pp. 999–1033.

Rust, J. (1994), 'Structural estimation of markov decision processes', *Handbook of econometrics* **4**, 3081–3143.

Schultz, T.W. (1961), 'Investment in human capital', *The American Economic Review* **51**(1), 1–17.

Sheshinski, E. (1967), 'Tests of the" learning by doing" hypothesis', *The review of Economics and Statistics* **49**(4), 568–578.

Song, Inseong and Pradeep K Chintagunta (2003), 'A micromodel of new product adoption with heterogeneous and forward-looking consumers: Application to the digital camera category', *Quantitative Marketing and Economics* **1**(4), 371–407.

A Sketch of Proof for Lemma 1

Proof. Under Assumptions 1 and 2, it is well known that the dynamic optimization problem has the associated Bellman equation (12). First define I_{ib} as the expected maximum current and future utility (or the *inclusive value*), *net of* the instantaneous purchase utility $u_{ib}^{I}+(1-\lambda)\varepsilon_{ibt}^{I}$, from making the purchase decision $B_{it}=b$:

$$
\begin{aligned}
I_{ib} &\equiv I_i\left(B_{it},\mathbf{S_t}\right)\\
&= \mathbb{E}\left[\max_{D_{it}} U_i\left(B_{it},D_{it},\mathbf{S_t}\right)|B_{it},\mathbf{S_t}\right]-u_{ib}^{I}-(1-\lambda)\,\varepsilon_{ibt}^{I} \qquad (19)
\end{aligned}
$$

where $u_{ib}^{I}\equiv u_i^{I}\left(B_{it},C_{it-1},\mathbf{X_t},\mathbf{P_t}\right)$. I_{ib} can be roughly considered as the inclusive value for *investing* in a given camera.

We now restrict the parameter in the production side, i.e. $\gamma_j=\overline{\gamma}_k$ if $j\in k$, and $\sigma_{kk'}$ captures the switching cost. With this restriction in the parameter space, we can rewrite the instantaneous utility from taking pictures, as

$$
\begin{aligned}
u_{ibd}^{II} &= u_i^{II}\left(B_{it},D_{it},C_{it-1},V_{it}\right)\\
&= D_{it}\left\{\alpha-\overline{\gamma}_k\left(1-V_{it}-\sigma_{kk'}^2-\sigma_{\omega}^2\right)+\kappa\right\}\\
&= \overline{u}_{ikd}^{II}.
\end{aligned}
$$

That is, the restriction in parameter space partially restricts the state space. Further more, combine (12) and (19) to rewrite the inclusive value of picture taking net of the purchase utility:

$$
\begin{aligned}
I_i\left(B_{it},\mathbf{S_t}\right) &= \mathbb{E}\left[\max_{D_{it}} U_i\left(B_{it},D_{it},\mathbf{S_t}\right)|B_{it},\mathbf{S_t}\right]-u_{ib}^{I}-(1-\lambda)\,\varepsilon_{ibt}^{I}\\
&= u_i^{II}\left(B_{it},D_{it},C_{it-1},V_{it}\right)+\lambda\varepsilon_{idt}^{II}+\\
&\quad \delta\ \mathbb{E}\left[\max_{B_{it+1},D_{it+1}} U_i\left(B_{it+1},D_{it+1},\mathbf{S_{t+1}}\right)|B_{it},D_{it},\mathbf{S_t}\right]\\
&= \overline{u}_{ikd}^{II}+\lambda\varepsilon_{idt}^{II}+\delta\ \mathbb{E}\left[\max_{B_{it+1},D_{it+1}} U_i\left(B_{it+1},D_{it+1},\mathbf{S_{t+1}}\right)|B_{it},D_{it},\mathbf{S_t}\right].
\end{aligned}
$$

Now impose Assumption 3, i.e. the choice specific utility shock ε_{ibt}^{I} can be decomposed into two

parts:

$$\varepsilon_{ibt}^{I} = \varepsilon_{ikt}^{i} + \varepsilon_{ijt}^{ii} j_k$$

with the first part only relevant to decision to purchase a type of camera. Given this assumption, and take operator $\max_{B_{it} j \bar{B}_{it}, D_{it}} (.)$ on the original Bellman equation. We then have:

$$\begin{aligned}
\overline{U}_i \left(\overline{B}_{it}, D_{it}, V_{it}, \overline{C}_{it\ \ 1}, \Gamma_{it}\right) &= \mathbb{E}\left[\max_{B_{it} j \bar{B}_{it}, D_{it}} U_i \left(B_{it}, D_{it}, \mathbf{S_t}\right)\right] \\
&= \max_{B_{it} j \bar{B}_{it}, D_{it}} \left\{ u_i^I \left(B_{it}, C_{it\ \ 1}, \mathbf{X_t}, \mathbf{P_t}\right) + (1 \quad \lambda)\, \varepsilon_{ibt}^{I} + I_i \left(B_{it}, \mathbf{S_t}\right) \right\} \\
&= \theta \log \left(\overline{P}_k\right) \ \ \mathbf{1}(k \quad 1) + (1 \quad \lambda)\, \varepsilon_{ikt}^{i} + \mathbb{E}\left[\max_{B_{it} jk} \sum_{j2k} \left\{ \theta \log \left(\frac{P_{jt}}{\overline{P}_k}\right) \ \ \mathbf{1}\left(B_{it} = j\right) + \right.\right. \\
&\qquad \left.\left. X_j \beta_i \ \ \mathbf{1}\left(C_{it} = j\right) + (1 \quad \lambda)\, \varepsilon_{ijt}^{ii} j_k \right\} jk, C_{it}, \mathbf{X_t}, \mathbf{P_t} \right] + I_i \left(B_{it}, \mathbf{S_t}\right) \\
&= \theta \log \left(\overline{P}_k\right) \ \ \mathbf{1}(k \quad 1) + \overline{u}_{ikd}^{II} + (1 \quad \lambda)\, \varepsilon_{ikt}^{i} + \lambda \varepsilon_{idt}^{II} + \Gamma_{ik} + \\
&\qquad \delta \ \ \mathbb{E}\left[\max_{\overline{B}_{it+1}, D_{it+1}} \left\{ \max_{B_{it} j \overline{B}_{it}, D_{it}} U_i \left(B_{it+1}, D_{it+1}, \mathbf{S_{t+1}}\right) \right\} j B_{it}, D_{it}, \mathbf{S_t} \right] \\
&= \theta \log \left(\overline{P}_k\right) \ \ \mathbf{1}(k \quad 1) + \overline{u}_{ikd}^{II} + (1 \quad \lambda)\, \varepsilon_{ikt}^{i} + \lambda \varepsilon_{idt}^{II} + \Gamma_{ik} + \\
&\qquad \delta \ \ \mathbb{E}\left[\max_{\overline{B}_{it+1}, D_{it+1}} \overline{U}_i \left(\overline{B}_{it+1}, D_{it+1}, V_{it+1}, \overline{C}_{it}, \Gamma_{it+1}\right) j B_{it}, D_{it}, \mathbf{S_t} \right] \qquad (20)
\end{aligned}$$

where

$$\Gamma_{ik} \left(C_{it\ \ 1}, \mathbf{X_t}, \mathbf{P_t}\right) = \mathbb{E}\left[\max_{B_{it} jk} \sum_{j2k} \left\{ \theta \log \left(\frac{P_{jt}}{P_k}\right) \ \ \mathbf{1}\left(B_{it} = j\right) + X_j \beta_i \ \ \mathbf{1}\left(C_{it} = j\right) + (1 \quad \lambda)\, \varepsilon_{ijt}^{ii} j_k \right\} jk, C_{it\ \ 1}, \mathbf{X_t}, \mathbf{P_t} \right]$$

denotes the expected maximum of static utility from camera purchasing, given that the consumer decides to purchase a camera in category $k = 0, 1, 2$, and $\overline{P}_k$ is the time invariant, category-specific average price.

Note that if the historical dependence pattern of Γ_{ik} is restricted, Equation (20) will define a new fixed point on a restricted state space. To see this, impose that Γ_{ik} is first order Markov, hence $\overline{B}_{it}$, $\overline{C}_{it\ \ 1}\, D_{it}$ and Γ_{it} is sufficient statistics for B_{it}, D_{it}, $\mathbf{S}_t$, and

$$\begin{aligned}
\overline{U}_i \left(\overline{B}_{it}, D_{it}, V_{it}, \overline{C}_{it\ \ 1}, \Gamma_{it}\right) &= \theta \log \left(\overline{P}_k\right) \ \ \mathbf{1}(k \quad 1) + \bar{u}_{ikd}^{II} + (1 \quad \lambda)\, \varepsilon_{ikt}^{i} + \lambda \varepsilon_{idt}^{II} + \Gamma_{ikt} + \\
&\qquad \delta \ \ \mathbb{E}\left[\max_{\bar{B}_{t+1}, D_{it+1}} \overline{U}_i \left(\overline{B}_{it+1}, D_{it+1}, V_{it+1}, \overline{C}_{it}, \Gamma_{it+1}\right) j B_{it}, V_{it}, \overline{C}_{it\ \ 1}, \Gamma_{it} \right].
\end{aligned}$$

is the new Bellman equation. □

B Model Specification for the Reduced Form Evidence

B.1 Proxy for Picture Quality

We take the residual from regressing log number of views per unit time, $\log(views_{pit})$, for picture p taken by i at t, on a set of co-variates controlling for fixed effects. Specifically, we control for i) the individual fixed effects,[32] ii) the length of time when a picture is in display, iii) number of pictures taken in the same month; iv) upload activities of the individual when the picture is in display, and v) the number of views of the 50 adjacent pictures.

This said, we use a simple reduced form model to project picture quality. For picture l photographed in period t by individual i:

$$\begin{aligned}\log(views_{pit}) &= \sum_{\tau}\lambda_{\tau}^{d}\ 1(\#months\,displayed_{it}=\tau)+\sum_{\tau}\sum_{m=0}^{4}\lambda_{m,\tau}^{f}\ 1(\#uploads_{i,t+m}=\tau)\\ &\quad+\sum_{k=\ 1}^{25}\lambda^{v}{}_{k}views_{p+k,it}+\alpha_i+\zeta_{pit}\\ &\quad X_{pit}\Lambda+\alpha_i+\zeta_{lt}\end{aligned} \tag{21}$$

where X_{lt} denotes the set of control variables, while α_i denotes the individual fixed effect. Note that we only control for individual fixed effects in Equation (21), and the effects of cameras (physical capital) is captured in the residual, ζ_{lt}.

We obtain the within-individual estimates $\hat{\Lambda}$, and then project the individual fixed effects $\hat{\alpha}_i$, and the errors $\hat{\zeta}_{lt}$. The proxy for picture quality is then:

$$q_{pit} \quad \hat{\alpha}_i+\hat{\zeta}_{pit}. \tag{22}$$

And its within-individual-period average

$$\bar{q}_{it}=\hat{\alpha}_i+\sum_{pji,t}\hat{\zeta}_{pit}$$

[32] Leaving camera fixed effect in the residual, since we need camera fixed effects in the final proxy of quality.

will enter the structural estimation as $quality_{it}$.

B.2 Learning By Doing

In order to demonstrate the presence of learning by doing, in Section 3 I estimate the following linear equation:

$$\log(views_{pit}) = \lambda_1^e expr_{it} + \lambda_2^e expr_{it}^2 + \lambda_1^p \sum_{\tau<t} pic_{i\tau} + \lambda_2^p \left(\sum_{\tau<t} pic_{i\tau}\right)^2 + X_{pit}\Lambda + \alpha_i + c_{it} + \nu_{pit}. \quad (23)$$

where $\sum_{\tau<t} pic_{i\tau}$ denotes total number of pictures taken prior to period t, α_i denotes individual fixed effect, and $c_{it} \quad \sum_j \delta_j \; 1(cam_{it} = j)$ denotes camera fixed effect. In the first and second column of Table 4 presents regression results with and without the total number of pictures as co-variates, respectively.

C Additional Tables and Figures

C.1 Tenure and the Number of Views

Table 10: Tenure and the Number of Views (Robustness Checks)

	$\log(Views/Months)$	$\log(Views/Months)$
Log cum. views for pics taken in the previous month	0.297***	0.294***
	(0.001)	(0.001)
Months in Flickr (00s)	0.673***	-0.019
	(0.015)	(0.021)
Months squared (0000s)	-0.079***	0.183***
	(0.009)	(0.010)
Months spent taking pictures (00s)		1.322***
		(0.029)
Months taking pictures squared (0000s)		-1.051***
		(0.020)
Display month dummies	Yes	Yes
Views of adjacent pics	Yes	Yes
Pictures taken now	Yes	Yes
Pictures uploaded now and subsequently	Yes	Yes
Individual and camera fixed effects	Yes	Yes
Rsq.	0.158	0.160
obs.	1907415	1907415

Notes:

1. Appendix B.2 documents model specification in detail.

2. This robustness check exercise adds lag monthly views to control for serial correlation in the error term, which might confound with learning by doing.

2. The sample is restricted to the pictures shot by digital SLR and compact cameras; we exclude those shot in film since we do not observe the exact camera model, and those shot by cellphones since some smart-phones have photo-sharing software installed that can potentially plague our estimates.

3. Total number of pictures taken in the past (and its quadratic term) are in the unit of *thousands*.