Stuck in the Adoption Funnel: The Effect of Delays in the Adoption Process on Ultimate Adoption *

Anja Lambrecht
London Business School
alambrecht@london.edu

Katja Seim
University of Pennsylvania
kseim@wharton.upenn.edu

Catherine Tucker
MIT Sloan School of Management
cetucker@mit.edu

Abstract

Many firms have introduced internet-based customer self-service applications, such as online payments or brokerage services. Despite high signup rates, these services have generally been unsuccessful in convincing customers to move most of their dealings with the firm online; few customers use them intensively. We investigate whether the multi-stage nature of the adoption process (an 'adoption funnel') for such technologies can explain this low take-up. We develop an empirical methodology to estimate a multi-stage adoption process within a discrete-time hazard framework with unobserved heterogeneity. We find that delays in earlier stages of the adoption process reduce a customer's probability of moving to intensive usage. Our results suggest significant cost-saving opportunities from reducing delays in the adoption funnel.

Keywords: Online Banking, Technology Adoption, Adoption Process, Adoption Funnel, Online Security, Self-Service Technology.

*We gratefully acknowledge financial support from the NET Institute, the Wharton Mack Center for Technological Innovation, and the Deutsche Forschungsgemeinschaft. We thank Michael Brown, Avi Goldfarb, Bruce Hardie, Jan-Stefan Lambrecht, Nicole Montgomery, John Roberts, Bernd Skiera, Catalina Stefanescu, Nader Tavassoli, Raphael Thomadsen, Naufel Vilcassim, Joel Waldfogel, Pai-Ling Yin, and seminar participants at the Rotterdam Research Day and at Frankfurt University for comments. We thank Jikyung Kim for research assistance. We are grateful to the bank, Acxiom Deutschland GmbH, and Hoppenstedt Firmeninformationen GmbH for sharing their data. An earlier version of this paper circulated under the title "Adoption and Usage of Online Services in the Presence of Complementary Offline Services: Retail Banking".

1 Introduction

Firms in many industries, such as banking, education and healthcare, allow customers to access and manage their own accounts online. Online self-service applications can save money and time for both firms and customers compared to traditional call centers or face-to-face interaction. They also offer opportunities for firms to improve customer service, target previously unserved customers, and cross-sell. However, research suggests that such online applications have not fully lived up to their productivity promises (Gordon (2000)). Despite widespread Internet diffusion, customers often fail to substantially use these services (Goldfarb and Prince (2007)). One notable feature of such technologies is that they require users to navigate a multi-stage adoption process. We explore whether shocks that delay a customer's momentum through this multi-stage process can help explain this 'usage gap'.

We investigate empirically the relationship between time spent in different stages of the adoption process and whether the customer ultimately uses the technology substantially. We use customer-level online banking data from a German retail bank on the timing of sign-up for the service, initial login into the service, and the initial online transaction, as well as the customers mix of online and offline transactions. These data allow us to follow customers through the adoption process and evaluate how delays affect the timing of progress and the eventual adoption outcome. We characterize this adoption process as an 'adoption funnel' to reflect the extent of customer attrition: Only 73% of customers who sign up for online banking ever log in; only 63% of customers ever complete an online transaction; and only 24% ever make more transactions online than offline. This usage gap matters. For online banking, there is evidence that the usage gap reduces profits, as even after controlling for self-selection active users of online banking generate larger revenues for the bank (Hitt and Frei (2002), Lambrecht (2005)).

Many customers take a long time to move to the next stage in the adoption funnel. The move from signup to first login takes on average more than a month. For 26% of users, the subsequent move from first login to the first online transaction does not occur in the same month. We explore whether slow progress through later stages of adoption is linked to how fast customers move through earlier stages.

Methodologically, we need to allow for the inter-relation of multiple stages of adoption. Our paper proposes an empirical framework that extends the typical discrete time hazard model to integrate all stages of a multi-stage adoption process into a single model. We disentangle the effect of early-stage delays on later-stage adoption from a customer's baseline probability to adopt and allow for unobserved heterogeneity in both the effect of delays and in a customer's baseline probability through latent class.

Our substantive contribution lies in the empirical result that the different stages of a multi-stage adoption process are indeed interrelated and should not be treated as separate: A delay in an early stage of the adoption process strongly reduces a customer's probability of substantially using the new technology. Importantly, this explains the previously observed usage gap as arising from the nature of the adoption process itself, going beyond explanations based on consumer heterogeneity. The knock-on effect of early delays, possibly caused by memory loss and the interruption of learning and momentum along the process, slows adoption and leads to the empirically observed gap between signup and usage.

We use our parameter estimates to predict the cost savings realized by the bank from customers passing through the adoption funnel faster. Reducing the time between customer signup and initial login significantly increases incremental cost savings from online banking by increasing the likelihood of moving to substantial usage. The same holds for reducing the delays between initial login and the first online transaction. This highlights how important it is for the bank to monitor and manage the customer-level adoption process.

Our results also highlight that the extensive regulatory efforts to ensure data security could potentially have unintended consequences for the diffusion of online services. In the case we study, the bank initiated the multi-stage process in response to EU directive 95/46/EC.[1] To

[1]Directive 95/46/EC covers the protection of individuals with regard to the processing of personal data and the free movement of such data.

satisfy the opaque conditions for 'electronic consent,' the bank chose to use traditional written means to prove that the consumers had consented to their data being made accessible online. In response to pressure from national regulators, German banks further developed a system of transaction authentication numbers (TANs) for each online transaction, which give customers another hoop to jump through when first making a transaction online.[2] Our research emphasizes that multi-stage authentication processes are more vulnerable to delays and may lead to lower levels of adoption.

The paper is organized as follows. Section 2 explains how our findings contribute to the previous literature. Section 3 presents our data. Section 4 records the extent of customer attrition along the adoption funnel for this technology. Section 5 discusses our model. Section 6 contains the results of our estimation. Section 7 investigates the cost implications of delays in the adoption process. Section 8 summarizes our results and discusses how they apply to other customer self-service technologies.

2 Literature Review

We ask whether delays in the multi-stage adoption process for online technologies can explain slow diffusion. Our work draws upon two literatures: First, research on diffusion and adoption of new technologies, in particular where adoption involves multiple stages; and second, work from psychology and consumer behavior on memory, learning and momentum.

Starting with Bass (1969), substantial research focuses on the diffusion of innovations across individuals (see, among others, Van Den Bulte and Stremersch (2004)) or across countries (Tellis et al. (2003)). These aggregate diffusion studies usually treat the outcome of the individual adoption decision as a single discrete choice. As pointed out by Rogers (2003), however, the adoption process often requires the completion of several distinct stages involving multiple decision-makers

[2]Security concerns like this are not limited to the German market. In the US, the Federal Financial Institutions Examination Council recommended in 2006 that banks use multi-factor identification for authentication as opposed to the commonly used username/password combination.

or other complicating factors.

Our research fills a gap in the empirical literature on technology adoption, which while recognizing that the adoption process is important, has not had the data to model it with sufficient accuracy before. Kalish (1985), who distinguishes between awareness and adoption for innovations, for example, lacks information on whether individuals are aware of an innovation and thus focuses on the adoption part of his model in estimation. Two solutions in the literature address these data difficulties: ex-post survey data (Beal et al., 1957) and modeling assumptions about which exogenous shifters affect each stage (Van Den Bulte and Lilien, 2007). We are able to circumvent the data difficulties that such techniques address because we have data on the different stages of the adoption process. This allows us to build on this literature's insight that it is important to separately model the different stages of the adoption process, and to model how the different stages of the adoption process interact with each other.

Our empirical analysis is unusual, in that we use data that include outcome measures for four different stages in the adoption process for online banking. This allows to apply the insights of a newer literature in marketing that emphasizes the flexibility of the discrete-time hazard model. For example, Seetharaman and Chintagunta (2003) demonstrate the model's usefulness for modeling unobserved shopping trips, and Van Den Bulte and Lilien (2007) show the model's flexibility in allowing for initial consideration in technology adoption. We extend the previous work by empirically estimating multiple stages of adoption in one estimation process.

Our multi-stage adoption model can address previous findings that many customers sign up for, but never use, self-service technologies. For example, Goettler and Clay (2006) report that approximately 40% of customers of an online grocer never place an order. Our results provide an additional explanation for this usage gap to those currently proposed in the literature, which emphasizes the degree of a technology's 'usefulness' (Davis (1989)) or customer heterogeneity (Meuter et al. (2005), Goldfarb and Prince (2007)).

Our finding that customers are less likely to adopt if their adoption process is delayed builds on behavioral research that has explored in other consumer-choice contexts how delays affect consumer behavior. First, we draw on research on how interruptions and delays negatively affect memory for tasks (Bahrick (1979), Speier et al. (1999) and Speier et al. (2003)). These laboratory studies find that delays generally damage recall of information. This suggests that in real-world environments, delays in an adoption process may possibly damage customer recall, making it harder to move to a later stage. Second, Soman (2003) suggests that customers tend to evaluate experiences more negatively as time goes on in a broad variety of contexts. For example, a customer may appreciate his dinner in a restaurant less after a long wait. Applied to technology adoption, their findings could imply that after a delay a customer may evaluate her first experience with a self-service technology as more negative than she did during the service encounter. This may lower her motivation to complete the adoption process and her assessment of the value of the service, resulting in attrition. Third, Dhar et al. (2007) find evidence of 'shopping momentum', as the likelihood of purchasing a particular product is higher for shoppers who previously purchased another, unrelated product. However, this momentum can be interrupted if customers are, for example, required to make payments from different payment accounts. Applying this concept of momentum to the adoption process implies that delays to the adoption process could cause a similar interruption of momentum, preventing progress to a later adoption stage.

3 Data and the Online Banking Industry

Our data come from confidential customer records from a major German retail bank over 23 months from September 2001 to July 2003. The bank introduced online banking in 1997. Its online service allows customers not only to monitor their checking, brokerage, and credit card accounts online, but also to initiate domestic and foreign wire transfers, the purchase or sale of brokerage account holdings, and the setup of recurring payments. Customers benefit from online banking because it is quicker to initiate a transaction online than in a branch. There are also cost savings for the bank. German estimates put the savings from processing an online-initiated wire transfer instead of a paper-based wire transfer, which is a significantly more popular transaction type in Europe than checks, at €0.50 - 1.00 (Wuebker and Hardock (2002) and conversations with the bank).

For each of 55,513 customers, the data include the date of signup for online banking, the monthly number of logins, the monthly number of online transactions broken down by type, and the monthly number of offline transactions. Offline transactions are not available by type, but include common cash withdrawals and check transactions. The bank counts a recurring transaction as online or offline depending on which channel the customer used to set it up. The data also include information on the age and gender of the primary account holder[3] and whether a customer has a brokerage account in addition to a checking account. We do not have information about other possible drivers of signup such as the type of transactions a customer usually conducts (for example the number of recurrent transactions).

Of the 55,513 customers, 3,200 had signed up for online banking prior to our data period. We use data on the 3,642 customers who signed up for the service during the 23-month span of the data for estimation and out-of-sample validation. Our estimates are therefore representative of the actions of the 'early majority' of adopters (Rogers (2003)). The average customer is 35.8 years old. 52% of customers are male. The customers are located in all 16 German states.

Another important influence for moving to the next stage in the adoption process is the amount of time customers have to devote to adoption and learning. To capture this we use panel data on the number of state sanctioned vacations and public holidays in each month. During school vacations, which vary state-by-state in their timing and length, many Germans leave their homes to travel. Germans, on average, take 26 vacation days a year (out of the average 25-30 day allowance), compared to the 11 days taken by Americans,[4] and go on roughly 1.6 vacation trips per year.[5] Vacation days affect a customer's ability to engage in the adoption process

[3]We do not observe whether the account is a joint account, but assume instead that the primary account holder manages the household's banking activities.

[4]Expedia.com, '2007 International Vacation Deprivation Survey Results', `http://www.expedia.com/vacationdeprivation` (accessed 08/16/07).

[5]Axel Springer Marketing Anzeigen, 'Tourismus 2002', `http://www.mediapilot.de/cda` (accessed 08/16/07).

because they usually separate a customer from her computer. A customer who leaves her home for vacation is unlikely either to receive the mailing of her online banking details, such as pre-assigned transaction-specific authorization numbers, or to carry these with her. Government agencies also warn consumers against using online banking on public computers at their vacation destination for security reasons.[6]

There is also variation in the number of state-level public holidays. Travel is less frequent on these holidays than at the beginning or end of school vacations but more frequent than during other periods of the year.[7] Spending public holidays at home gives consumers time to catch up on chores such as experimenting with online banking. The average effect of public holidays on the propensity to use online banking is therefore less clear than the effect of vacations. Since public holidays alter consumers' availability of time relative to a normal working day, we would expect the monthly probabilities of progressing through the funnel to differ across otherwise identical customers with different numbers of public holidays available to them during the month. We provide more detail on the correlates and distribution of vacations and public holidays in section A.3 in the Appendix.

4 The Adoption Funnel

4.1 The Four Stages of the Adoption Funnel

'Adoption' frequently refers to the consumer's decision to purchase or begin using a product or service. For customer self-service technologies, it is difficult to identify one discrete decision that indicates adoption. In the case of online banking, the customer goes through four successive stages before the bank realizes cost savings. We call this adoption process the 'Adoption Funnel.' The four stages are:

[6] Bundesamt für Sicherheit in der Informationstechnik, 'Brennpunkt: IT-Betrueger machen keinen Urlaub', `http://www.bsi-fuer-buerger.de/brennpunkt/urlaub.htm` (accessed 08/16/07).

[7] See `www.adac.de/Verkehr/Staukalender/default.asp` (accessed 08/16/07).

Table 1: Description of Variables used in Estimation

Variable Label	Variable Description	Mean	Std. Dev.
	Outcome Variables		
Customers who reach stage 1	Share ever evaluate (login)	0.730	0.444
Customers who reach stage 2	Share ever try (conduct transaction)	0.625	0.484
Customers who reach stage 3	Share ever use substantially (more than 50% of transactions)	0.237	0.426
	Endogenous Delay Variables		
Delay between Signup and Login	Months between signup and first login	1.219	2.262
Delay between Login and First Transaction	Months between first login and first transaction	0.687	1.871
	Explanatory Variables		
Male	Indicator variable for whether primary account holder is male	0.508	0.5
Age	Age of primary account holder	36.261	13.61
Brokerage	Whether the customer has a brokerage account	0.315	0.464
Vacation Days	Number of vacation days in state of residence of primary account holder in that month	4.435	5.515
Public Holidays	Number of public holidays in state of residence of primary account holder in that Month	0.903	0.976

n=3,642 customers who sign up for online banking between 9/2001 and 7/2003.

1. Signup: A customer signs up for a new internet service.
2. Evaluation: A customer logs into the service for the first time.
3. Trial: A customer processes the first transaction using the service.
4. Substantial usage: A customer processes the majority of their transactions using the service.

In our case, the first stage occurs when a customer submits a paper form to a bank branch indicating the wish to use online banking. The bank then sends back a paper letter detailing log-in identification. This system allows the bank to verify the users identity and securely transmit the login details. The creation of the login details is automated and takes between one and two working days. The customer is ready to log into the platform within five to seven working days after the bank's receipt of her application.

In the second stage the customer logs into the web site, possibly to check her account details online and investigate the functionality of the new service. We define the third stage, trial, as the customer initiating her first online transaction. Though evaluation (first login) and trial (first transaction) are often used synonymously, in our setting they are separated both functionally and empirically. The bank e-mails a customer her login details, which separately only allow for browsing of the online banking site. Two to three days later, the customer receives a list of 6-digit TANs. An online transaction can only be executed if it is verified with a TAN. Each TAN is only valid for a single transaction and cannot be reused. The process therefore imposes a short time lag between evaluation and trial. In addition, trial requires the customer to jump through an extra security hoop and thus the customer's interaction with the online banking site is more complex than in the evaluation stage.

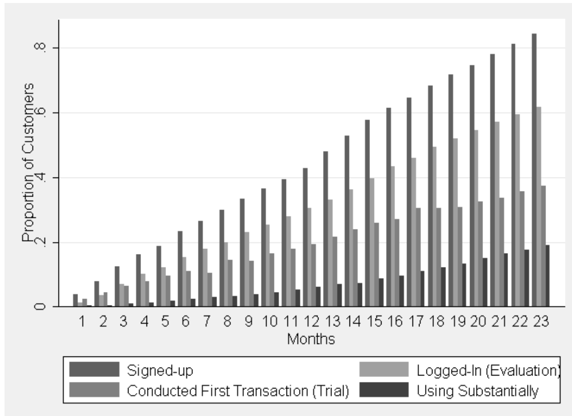
We find that 26% of customers do not conduct their first transaction in the month that they first login. Gaps like this between evaluation (first login) and trial (first transaction) are common. 95% of French and Italian online banking customers confirm that they have checked their account balance online in the past three months, while only 20-25% transferred money or paid bills online within the time period.[8]

There are many ways of defining substantial or intense usage. We are limited in defining a baseline because our data on total off-line transactions include the number of recurring transactions and cash withdrawals. We approximate 'substantial usage' to have been reached in the first month in which the customer conducts 50% of all recorded transactions in that month online. After attaining substantial usage based on this definition, the average ratio of online to total transactions is 0.63 for the remaining months in our data. This suggests that customers retain their status as substantial users beyond the initial transition. As a robustness check, we also check an alternative definition that defines substantial usage as 40% of transactions, with similar results (see Table A-5 in the Appendix).

While our empirical measures of the adoption funnel are developed in the context of online banking, the general framework applies to other self-service technologies. Depending on the service, the exact sequence of stages may differ or the stages may be more nuanced.

[8]Ensor, B.'French online banking forecast: 2008 to 2013.' Technical report, Forrester, February 2008, and Ensor, B. and A. Hesse 'Italian online banking forecast: 2008 to 2013.' Technical report, Forrester, February 2008.

Figure 1: Proportion of Customers in Different Stages of the Funnel over Time



4.2 Conversion along the Funnel

Figure 1 shows that the diffusion process differs considerably for the four stages. Significantly fewer customers get to the stage of substantial usage than sign up for online banking. Table 1 reports summary statistics for our data. Of all customers who signed up for the service, 73% logged in at least once and 63% completed at least one transaction online. Only 24% exhibit substantial usage of online banking.

Figures 2 and 3 display the distributions of the number of months between signup and login and the number of months between the first login and the first transaction, which are our measures of delay. 55 percent of customers who go on to log in do not log in in the same month that they signed up. 26 percent of customers who ultimately go on to make a transaction do not conduct their first transaction in the same month they first login. On average (conditional on reaching that stage), customers spend 1.3 months between signup and login, 0.7 months between login and the first transaction, and 2.9 months between the first transaction and substantial usage. These estimates are biased downwards because we only observe the length of a particular stage if a customer reached the next stage within the two years of our panel. We adjust our discrete time hazard specification to account for the consequent censoring of the hazard rate.

Figure 2: Months between Signup and Login.

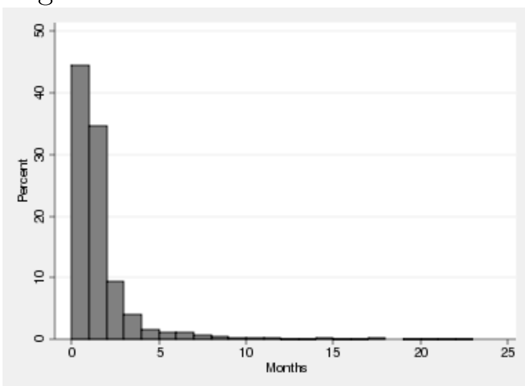
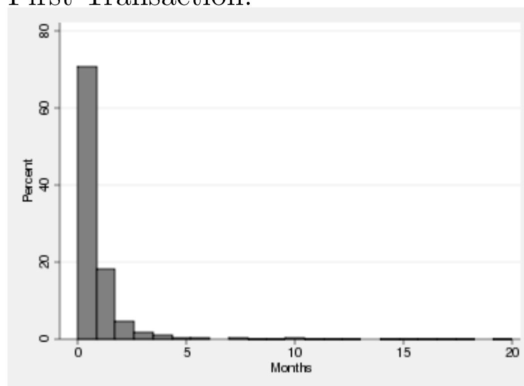


Figure 3: Months between Login and First Transaction.



5 Model of Timing of Multiple Adoption Decisions

It is empirically challenging to identify the effect of time spent in the adoption funnel on attrition, because unobserved customer heterogeneity may drive both current stage adoption and the length of the adoption process in previous stages. A positive correlation could merely reflect the fact that some customers are faster at adopting technology than others. We want, however, to explicitly model the relationship between a customer spending a long time between signup and login, and between login and the first transaction, and the likelihood of substantial usage.

We propose an adoption model that augments the standard discrete time hazard model to allow for multiple interrelated stages. The discrete time hazard model is an attractive alternative to continuous time hazard models for survival time data where only discrete time (for example the month of adoption) is available as opposed to precise timing (Allison (1982), Jenkins (1995)). Let grouping point t_m, $m = 1, ..., M = 23$, denote the end of month m, where $m = 1$ corresponds to September 2001. We observe data on the number of months elapsed between signup (event 0 at time t_{m^0}), first login (event 1 at time t_{m^1}), first transaction (event 2 at time t_{m^2}), and first substantial usage (event 3 at time t_{m^3}). We define the random variables D^T, $T = 1, 2, 3$, as

$D^1 =$ number of months customer spends between signup and first login $= (t_{m^1} - t_{m^0})$

$D^2 =$ number of months customer spends between first login and first transaction $= (t_{m^2} - t_{m^1})$

$D^3 =$ number of months customer spends between first transaction and substantial usage $= (t_{m^3} - t_{m^2})$

We specify the likelihood of observing a particular duration realization for D^T as a discrete time hazard model. The likelihood of individual i's transition time is then the product of the survivor function in each period that individual i does not transition and the hazard at the time of transition. The discrete time hazard function is given by

$$\begin{aligned}
\lambda_i^T(t_m|z_i(t_m), d_i^{T-1}, \beta_j^T) &= \Pr[t_{m-1} \leq D_i^T < t_m|z_i(t_m), d_i^{T-1}, \beta_j^T] \qquad (1)\\
&= \begin{cases} F\left(\beta_0^T(t_m - t_{m^{T-1}}) + \beta_j^T + z_i(t_m)'\beta_z^T + \beta_{dij}^T d_i^{T-1}\right) & T = 2, 3 \\ F\left(\beta_0^T(t_m - t_{m^{T-1}}) + \beta_j^T + z_i(t_m)'\beta_z^T\right) & T = 1 \end{cases}
\end{aligned}$$

with an associated discrete time survivor function

$$\begin{aligned}
S_i^T(t_m|z_i(t_{m^{T-1}}, ..., t_{m-1}), d_i^{T-1}, \beta_j^T) &= \Pr[D_i^T \geq t_m|z_i(t_{m^{T-1}}, ..., t_{m-1}), d_i^{T-1}, \beta_j^T] \qquad (2)\\
&= \prod_{s=m^{T-1}}^{m-1} \left(1 - \lambda_i^T(t_s|z_i(t_s), d_i^{T-1}, \beta_j^T)\right)
\end{aligned}$$

In equations (1) and (2) F refers to the distribution function of an unobservable *i.i.d.* shock to the hazard. We assume that this shock is drawn from a type 1 extreme value distribution, implying a logit specification. We assume that these unobservable shocks are drawn from the same distribution across all stages.

We include a time-varying stage-specific base hazard $\beta_0^T(t_m - t_{m^{T-1}})$ as a linear function of the time that has elapsed since a customer has entered a stage.[9] This reflects that a customer's propensity to adopt may fall with the time she has spent in a stage because, for example, she is

[9]Estimating a semi-parametric and a non-parametric version did not improve model fit.

more likely to forget her login details as time passes. It also helps us to disentangle the causal effect of a delay in an earlier stage from time spent in the current stage.

The term β_j^T captures the customer's intrinsic unobserved value of adopting a new technology. In the estimation of duration models with a single spell per customer, such as a single adoption decision, it is econometrically not possible to identify individual-level unobserved heterogeneity alongside time-varying co-variates, such as a time-varying baseline.[10] Deviations from the average adoption probability can either serve to identify individual-level heterogeneity, making it impossible to identify variation across time or, alternatively, to identify variation across time, which in turn does not allow to identify individual-level unobserved heterogeneity (Allison and Christakis (2006)). Consequently, duration models often do not account for unobserved heterogeneity.[11] In continuous time duration models, Heckman and Singer (1984) have attempted to include unobserved heterogeneity by relying on distributional assumptions to identify heterogeneity (similarly, see Meyer (1990)), but document that estimates are highly sensitive to the choice of a particular distribution of the unobserved heterogeneity. We accommodate the conflicting demands by estimating a time-varying hazard and allowing for unobserved heterogeneity across finely defined, homogeneous customer groups, denoted as j. Customer groups are defined based on an individual's age and gender. To further avoid reliance on distributional assumptions that may lead to unstable estimates, our estimation of unobserved heterogeneity relies on latent class (Kamakura and Russell (1989)).

We further assume that the base hazard, β_j^T, varies across the stages of the adoption model reflecting the stages' differing average levels of adoption. The latent class specification introduces correlation between a customer's base hazards in different stages. Hence, unobserved heterogeneity that leads to delays in earlier stages also affects adoption in later stages where delays enter as

[10] The problem is similar if, as in our setting, there are multiple adoption decisions but stage-specific intercepts and co-variates.

[11] Identification of individual-level heterogeneity alongside time-varying co-variates is possible with multiple spell data, such as is used in models for repeat purchase behavior, see for example Seetharaman and Chintagunta (2003) and Gupta (1991). To ensure stable estimation Gupta (1991) limits his sample to households with at least 10 purchases over two years.

endogenous variables. By explicitly allowing for unobserved heterogeneity to affect both current stage adoption and the length of the previous stage's adoption process, we disentangle the causal effect of delays on adoption from unobserved heterogeneity.

We allow the hazard rate to vary with time-variant and time-invariant regressors, $z_i(t_m)$. β_z^T measures the effect of these characteristics on adoption. As time-invariant regressors, we include the observable customer attributes age, gender and whether or not a customer has a brokerage account. We also include time-variant shifters to adoption: the number of vacation days and public holidays during the month are potential exogenous sources of delays to moving through the adoption process. We exploit the fact that Germany uses a staggered system of school vacations and public holidays across states, leading to wide variation in the timing of states' school vacations for different customers in our data set. The decreased accessibility of online banking while traveling implies that a customer who signs up for the service in a vacation-heavy month likely spends longer between sign-up and login than a customer who signs up at another time, for reasons unrelated to unobservable customer tastes for technology. Though in our setting the interruptions to the adoption process are vacations and holidays, it is plausible that customers respond similarly to delays that are caused by other exogenous factors as to delays that are caused by vacation and that our estimates are representative of the effect of delays that come from a greater variety of sources.

We use information on the gender and age of customers as a source of observable heterogeneity that might moderate the effect of delays on adoption. We define β_{dij}^T as a linear function of these characteristics:

$$\beta_{dij}^T = \gamma_{d0j}^T + \gamma_{d1}^T MALE_i + \gamma_{d2}^T AGE_i \tag{3}$$

We allow for unobserved heterogeneity in the effect of delays on adoption by group j, γ_{d0j}^T.

Our likelihood for stages 1-3, where i is the individual and j is their group is therefore,

$$
\begin{aligned}
l^1_{ij}(t_{m^1}) &= F(\beta^1_j + \beta^1_0(t_{m^1} - t_{m^0}) + z_i(t_{m^1})'\beta^1_z) \\
&\times \prod_{s=m^0,\ldots,m^1-1} \left(1 - F(\beta^1_j + \beta^1_0(t_s - t_{m^0}) + z_i(t_s)'\beta^1_z)\right) \\
l^2_{ij}(t_{m^2}) &= F(\beta^2_j + \beta^2_0(t_{m^2} - t_{m^1}) + z_i(t_{m^2})'\beta^2_z + \beta^2_{dij}d^1_{ij}) \\
&\times \prod_{s=m^1,\ldots,m^2-1} \left(1 - F(\beta^2_j + \beta^2_0(t_s - t_{m^1}) + z_i(t_s)'\beta^2_z + \beta^2_{dij}d^1_{ij})\right) \\
l^3_{ij}(t_{m^3}) &= F(\beta^3_j + \beta^3_0(t_{m^3} - t_{m^2}) + z_i(t_{m^3})'\beta^3_z + \beta^3_{dij}d^2_{ij}) \\
&\times \prod_{s=m^2,\ldots,m^3-1} \left(1 - F(\beta^3_j + \beta^3_0(t_s - t_{m^2}) + z_i(t_s)'\beta^3_z + \beta^3_{dij}d^2_{ij})\right)
\end{aligned} \tag{4}
$$

We observe considerable right-censoring corresponding to customers who do not complete the transition process by the end of our sample at $t = t_M$. Let δ^T_{ij} be an indicator variable that equals one if customer i of group j reaches stage T, $T = 1, 2, 3$. The likelihood that customer i does not complete the funnel is her survivor function for each period spent in the uncompleted stage, from the initial transition to the previous stage to the end of the sample period.

We use full-information maximum likelihood and specify the joint likelihood of customer i's transition decisions to move from signup to login to first transaction and ultimately to substantial usage. We obtain the unconditional joint likelihood of customer i's in group j's transition history as:

$$
\begin{aligned}
l_{ij} = \int &\big(l^1_{ij}(t_{m^1}|\beta_j, \beta_{dij})l^2_{ij}(t_{m^2}|d^1_{ij}, \beta_j, \beta_{dij})l^3_{ij}(t_{m^3}|d^1_{ij}, d^2_{ij}, \beta_j, \beta_{dij})\delta^3_{ij} \\
&+ l^1_{ij}(t_{m^1}|\beta_j, \beta_{dij})l^2_{ij}(t_{m^2}|d^1_{ij}, \beta_j, \beta_{dij})S^3_{ij}(t_M|d^1_{ij}, d^2_{ij}, \beta_j, \beta_{dij})\delta^2_{ij} \\
&+ l^1_{ij}(t_{m^1}|\beta_j, \beta_{dij})S^2_{ij}(t_M|d^1_{ij}, \beta_j, \beta_{dij})\delta^1_{ij} + S^1_{ij}(t_M|\beta_j, \beta_{dij})(1 - \delta^1_{ij})\big)\, dG(\beta_j, \beta_{dij})
\end{aligned} \tag{5}
$$

We use latent class techniques to non-parametrically represent unobserved customer heterogeneity in customers' intrinsic probabilities to adopt and in the effect of delays. We assume C

homogenous latent classes with relative sizes:

$$p_c = \exp(\theta_c)/\sum_c \exp(\theta_c) \tag{6}$$

Each customer's contribution to the likelihood is then obtained as the weighted average of the customers likelihood under the class-specific parameters:

$$\begin{aligned} l_{ij} \quad = \quad & \sum_c p_c \big(l^1_{ij}(t_{m^1}|\beta^c_j, \beta^c_{dij}) l^2_{ij}(t_{m^2}|d^1_{ij}, \beta^c_j, \beta^c_{dij}) l^3_{ij}(t_{m^3}|d^1_{ij}, d^2_{ij}, \beta^c_j, \beta^c_{dij}) \delta^3_{ij} \\ & + l^1_{ij}(t_{m^1}|\beta^c_j, \beta^c_{dij}) l^2_{ij}(t_{m^2}|d^1_{ij}, \beta^c_j, \beta^c_{dij}) S^3_{ij}(t_M|d^1_{ij}, d^2_{ij}, \beta^c_j, \beta^c_{dij}) \delta^2_{ij} \\ & + l^1_{ij}(t_{m^1}|\beta^c_j, \beta^c_{dij}) S^2_{ij}(t_M|d^1_{ij}, \beta^c_j, \beta^c_{dij}) \delta^1_{ij} + S^1_{ij}(t_M|\beta^c_j, \beta^c_{dij})(1 - \delta^1_{ij}) \big) \end{aligned} \tag{7}$$

6 Results

We split our data into an estimation sample and a holdout sample of equal sizes. We estimate our model based on all customers in the estimation sample that logged into the online banking site at least once. For customers that did not transition to login, we would not observe any delay between signup and login and hence cannot estimate an effect of such a delay on a transition to a later stage. Among the 1,297 customers of our estimation sample that logged in, 1,095 customers transitioned to a first transaction and 419 customers to substantial usage during the period of our data. Overall, our data include 2,849 observations in stage 1 (transition to first login), 3,888 observations in stage 2 (transition to first transaction) and 9,835 observations in stage 3 (transition to substantial usage). The large number of mainly non-adoption observations in stage 3 relative to the other two stages reflects that customers move only slowly to substantial usage.[12]

As laid out in section 5, we identify unobserved heterogeneity in customers' underlying

[12] As discussed by Allison (1982), the larger sample that results from ordering data to accommodate discrete hazard techniques does not bias standard errors downwards.

Table 2: Model Fit Comparison

Model Specification	Log-Likelihood	# Parameters	BIC
1 segment	-5170.1	33	10487.9
2 segments	-5155.2	39	10485.1
3 segments	-5152.5	45	10506.5
4 segments	-5150.2	51	10528.8

propensity to adopt technologies and in the effect of delays on adoption at the level of distinct customer groups. We group customers into 88 distinct groups based on the three individual-level observed characteristics gender (male vs. female), age (groups of 5 years of age) and state of residence. We then estimate the model laid out in section 5 and cluster standard errors at the group level using a bootstrap.

As we capture unobserved heterogeneity through latent class, we first compare model specifications with different numbers of segments based on the BIC. Table 2 shows that accounting for unobserved heterogeneity in the effect of delays on adoption significantly improves the estimation. The model with two latent classes best reflects unobserved heterogeneity in our data. Our two-segment model also dominates a specification that allows for unobserved heterogeneity on the intercept but not on delays (BIC of 10504), indicating that customers are indeed heterogeneous in their reaction to delays. The chosen specification includes a linear time-varying baseline, which we found preferable to a specification that allowed the effect of time spent to decline more flexibly by including separate indicator variables for the first through eighth period after the transition to the current stage, before declining linearly for subsequent periods (BIC of 10519).

Table 4 summarizes the results of the two-segment model. The adjusted McFadden R-square is 0.314, which is considered a good overall model fit. The heterogeneity of the intercept terms reflects unobserved heterogeneity in an individual's baseline probability to adopt. On average, the parameter estimates for class 1 have a larger probability than those for class 2 (73.8% vs. 26.2%). In line with this, an individual-level analysis reveals that for 86.4% of customers the posterior probability for class 1 is greater than the posterior probability for class 2.

The covariates in our model provide a plausible source of exogenous variation of interruptions that can lead to delays. Vacation days in the signup month significantly delay the first login.[13] On average, one day of vacation in the sign-up month reduces the probability of a first login in a given month by 0.38 percentage points. This effect of vacation is consistent with Germans leaving home during vacations and not conducting online banking abroad. A customer's base propensity to transition from first login to first transaction declines in the customer's age. It is lower for male customers and those who have both brokerage and checking accounts with the bank. The parameter estimates of time spent in stages 1, 2 and 3 also indicate that the probability to move to the next stage decreases with the time a customer has so far spent in the current stage.

Our key variables of interest are how delays affect progress towards a first transaction and towards substantial usage. Overall, we find strong evidence for a negative effect of earlier delays on adoption in later stages. For class 1 the results show a highly significant negative effect of the delay between signup and login on the transition to first transaction (stage 2). Similarly, a longer delay between login and the first transaction significantly decreases the probability to ever transition to substantial usage (stage 3). The delay estimates for the much smaller class 2 are less negative in stage 2 and positive in stage 3, indicating heterogeneity in the effect of delays across customers.

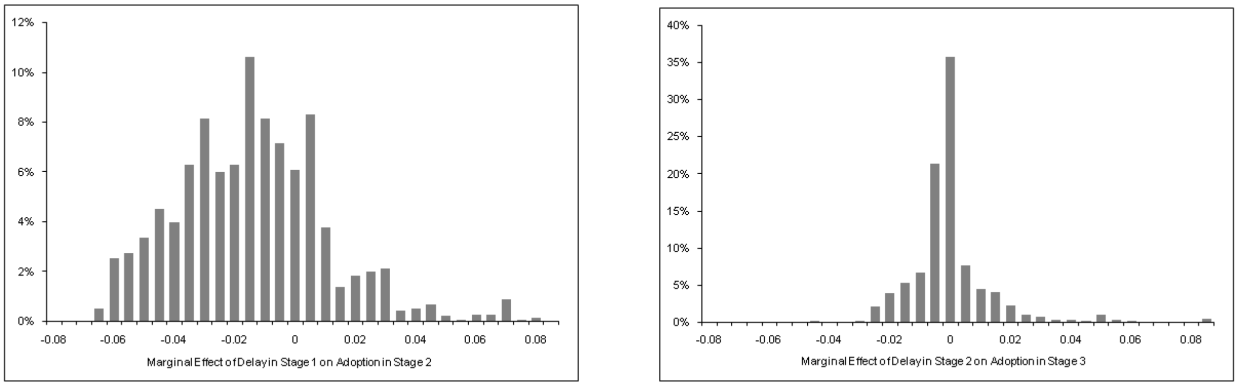
For the transition to a first transaction, age moderates the effect of delays, with younger customers experiencing more pronounced delays. Similarly, the negative effect of delays between signup and login on the transition to a first transaction is significantly less pronounced for men than for women. This is consistent with the finding that women on average have less online experience than men, and therefore have more to learn and retain about internet services.[14]

To better understand the individual-level effect of delays on adoption, we compute individual-level marginal effects based on the delay coefficients, the demographic interaction terms, and a

[13]We have also estimated specifications where we treat these vacation days and public holidays as exogenous instrumental variables in a reduced-form specification, and have conducted the standard over-identification/joint-significance tests. The results show that vacation days and public holidays are jointly significant in shifting delays. See page 33 in the appendix.

[14]Van Eimeren, B., H. Gerhard and B. Frees, 'ARD/ZDF-Online-Studie 2004: Internetverbreitung in Deutschland', Media Perspektiven, 8, 2004, 350-370.

Figure 4: Individual-level Marginal Effect of Delays



customer's posterior class probability. As the marginal effects vary across time periods, we aggregate a customer's monthly marginal effects to a customer-specific average. Our results confirm the importance of delays on an individual-level. The large majority of customers, 76.7%, have a negative marginal effect of a delay between signup and login on the transition to first transaction (Figure 4). We find similar results for the marginal effects of the delay between login and the first transaction on substantial usage. Here the marginal effect is negative for 76.1% of customers.

Breaking down these individual-level marginal effects based on observed demographics suggests that the effect of delays varies across demographic groups. Table 3 compares the median individual-level marginal effect of a delay to the median individual-level probability of transitioning to the first transaction and substantial usage by demographic group. We classify customers by gender and by their age relative to the median age in the estimation sample (35 years). The median overall effect of delaying a customer one month between signup and login (login and first transaction) reduces the probability of ever completing a first transaction (substantially using the technology) by 1.8 percentage points (0.4 percentage points), which compares to a baseline probability of 54.08% (5.12%). The marginal effect of a delay is thus significant in magnitude, in particular for delays between login and first transaction. It is for these later delays where heterogeneity across demographic groups is pronounced, with marginal effects being larger in

Table 3: Marginal Effects and Predicted Adoption Probabilities

	First Transaction		Substantial Usage	
	Marg. Eff.	**Pred. Prob.**	**Marg. Eff.**	**Pred. Prob.**
Male, <35	-0.0235	0.5416	-0.0033	0.0547
Male, ≥35	-0.0050	0.4900	-0.0021	0.0615
Female, <35	-0.0396	0.5977	-0.0075	0.0415
Female,≥35	-0.0304	0.5560	-0.0075	0.0468
All Customers	-0.0174	0.5408	-0.0037	0.0512

absolute value for female than male customers. Thus, the results support the interpretation that delays in the adoption funnel have a bigger effect on inexperienced customers with a greater need to learn, which is consistent with the previously cited laboratory findings on the impact of consumer memory on subsequent actions.

Our results have two important implications. First, they confirm that exogenous delays significantly affect whether and how fast customers move through the adoption process or get stuck in the adoption funnel. Second, our model results illustrate how a multi-stage adoption process can be treated in marketing, economics or management. Jointly estimating the full multi-stage process allows us to more accurately account for interactions between individual stages.

6.1 Validation

To ensure that our model results reflect the actual adoption behavior in our data, we first conduct an in-sample analysis and then turn to out-of-sample fit. The out-of-sample validation helps to demonstrate that the proposed model captures the underlying behavioral aspects of a multistage adoption process.

Figure 5 plots the stage-specific predicted and actual number of adopters over time. The predicted adoption decisions closely follow actual adoption behavior. This is supported by the low average and median percentage prediction errors in the range of -0.3% to 5.0% across all stages (part A in Table 5). Errors in later stages are slightly higher than in earlier stages, a result of early-stage prediction errors that feed through to later stages.

The RMSE indicates that the predicted number of adopters in stage 1 differs from the actual

Table 4: Main Results for Multi-Stage Adoption Process

Variable	Coefficient	Std. Error	Sig.	Marg. Eff.
Stage 1: Probability to transition from Sign-Up to Login				
Intercept stage 1 (class1)	0.182	0.146		0.044
Intercept stage 1 (class2)	-0.002	0.247		-4.5E-04
Male	0.063	0.110		0.015
Age	0.000	0.003		-4.8E-05
Brokerage account	0.113	0.077		0.027
Vacation Days in Month	-0.016	0.007	**	-0.004
Public Holidays in Month	-0.002	0.036		-0.001
Time spent in stage 1	-0.126	0.007	***	-0.030
Stage 2: Probability to transition from Login to First Transaction				
Delay between Signup and Login (class 1)	-0.416	0.130	***	-0.060
Delay between Signup and Login (class 2)	-0.165	0.304		-0.024
(Delay between Signup and Login) * Male	0.136	0.080	*	0.020
(Delay between Signup and Login) * Age	0.006	0.003	*	0.001
Intercept stage 2 (class1)	2.253	0.296	***	0.324
Intercept stage 2 (class2)	1.598	0.614	***	0.228
Male	-0.442	0.211	**	-0.063
Age	-0.021	0.008	***	-0.003
Brokerage account	-0.405	0.097	***	-0.058
Vacation Days in Month	-0.005	0.007		-0.001
Public Holidays in Month	-0.021	0.036		-0.003
Time spent in stage 2	-0.555	0.035	***	-0.133
Stage 3: Probability to Transition from First Transaction to Substantial Usage				
Delay between Login and First Transaction (class 1)	-0.437	0.219	**	-0.016
Delay between Login and First Transaction (class 2)	0.114	0.873		0.006
(Delay between Login and First Transaction) * Male	0.290	0.233		0.012
(Delay between Login and First Transaction) * Age	0.002	0.004		9.1E-05
Intercept stage 3 (class1)	-1.999	0.413	***	-0.074
Intercept stage 3 (class2)	-2.392	0.784	***	-0.130
Male	-0.085	0.326		-0.004
Age	0.006	0.008		2.4E-04
Brokerage account	-0.366	0.124	***	-0.015
Vacation Days in Month	0.009	0.008		3.8E-04
Public Holidays in Month	-0.034	0.042		-0.001
Time spent in stage 3	-0.171	0.022	***	-0.041
Stage Independent				
Parameter for class probability, θ_c	1.034	0.331	***	
Implied class probablility (class 1)	0.738			
Log-Likelihood	-5155.2			
Observations	1297			
McFadden R-Squared	0.317			
Adjusted McFadden R-Squared	0.312			

Bootstrapped Standard Errors, clustered at the Group Level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

number of adopters by 15.5 on average across all periods (stage 2: 15.4, stage 3: 16.9). This compares to 695.8 customers who have logged into online banking (574.4 customers who have done a first online transaction; 165.0 customers who have moved to substantial usage) at the end of a period on average across all 23 usage periods in our data. Thus, on average our prediction differs from the actual number of adopters by 2.2% in stage 1, 2.7% in stage 2 and 10.2% in stage 3. The slightly higher ratio in stage 3 is a result of the smaller sample size in that stage.

For out-of-sample testing, we select the customers with at least one login from our holdout

Figure 5: Predicted vs. Actual Number of Adopters Across Stages: Estimation Sample

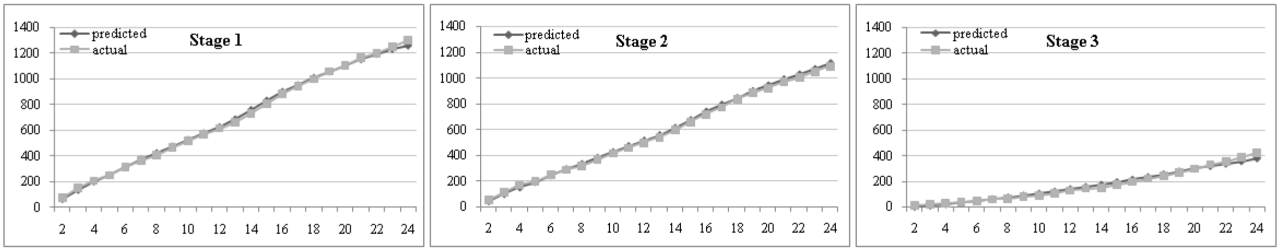
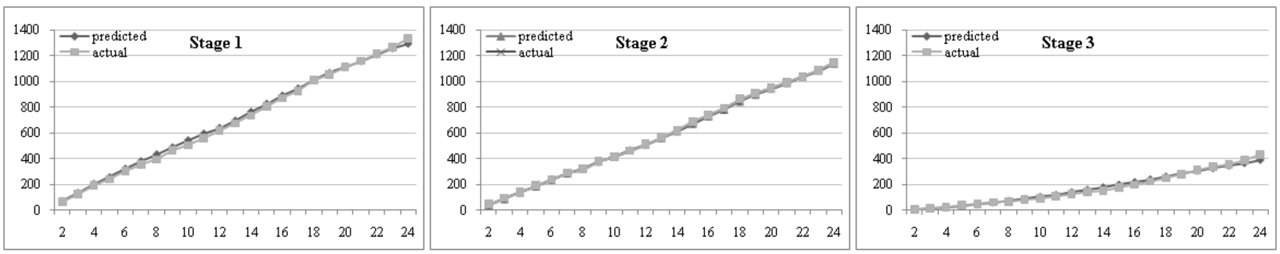


Figure 6: Predicted vs. Actual Number of Adopters Across Stages: Holdout Sample



sample (1,337 customers). Similar to our estimation sample, we plot the predicted versus actual number of adopters over time (Figure 6). Again, predicted adoption follows actual behavior. The percentage prediction errors and the RMSE (part B in Table 5) are similar to the estimation sample and indicate out-of-sample validity.

7 Implications

Banks invest in online banking to migrate paper-based transactions online and reduce the cost of processing transactions. However, one possible implication of our findings is that banks currently do not fully benefit from this technology, because customers' delays in progressing through the adoption funnel inhibit substantial usage. In this section we assess the impact of customers'

Table 5: Model Validation

	A: Estimation Sample			**B: Holdout Sample**		
	Stage 1	**Stage 2**	**Stage 3**	**Stage 1**	**Stage 2**	**Stage 3**
Average % prediction error	0.3%	-0.3%	3.1%	2.5%	-0.6%	0.9%
Median % prediction error	1.3%	2.1%	5.0%	2.5%	-0.3%	2.0%
RMSE	15.5	15.4	16.9	20.0	6.4	15.8
Average # of adopters by end of period	695.8	574.4	165.0	697.5	585.2	169.9
RMSE/average # of adopters	2.2%	2.7%	10.2%	2.9%	1.1%	9.3%

delays on firms' cost savings. We consider both an acceleration of the first stage (sign-up to login) and of the second stage (login to transaction) of the adoption funnel.

We do not model the source of customers' delays or how the bank can influence them (see below for a discussion of how a firm can influence delays). Instead we consider how decreases in delays at different stages of the funnel feed through the adoption process to affect the expected number of online transactions per customer and the bank's associated cost savings.

We use our individual-level parameter estimates to calculate the incremental savings of reducing delays by one month. We compute monthly adoption probabilities under actual delays for the period of our data and simulate adoption probabilities if delays were reduced by one month. We then predict cost savings that result from the altered adoption behavior as customers start conducting a substantial share of their transactions online sooner. We use a conservative estimate of cost-savings per online transaction of €0.50.

We discount the monthly savings to the first period in our data to get total individual-level incremental cost savings to the bank. This estimate indicates how much the bank could have saved if in the first period of our data it had set up a policy that shortened delays between signup and login and between login and the first transaction.

Our results show that banks can significantly increase their cost savings by helping customers move faster through the adoption funnel. Shortening the delay between signup and login by one month increases cost savings by 2.0% for customers who currently do not transition immediately. This result is a lower bound that underestimates the full effect because of conservative assumptions: First, it reflects only the knock-on effect of delays from an earlier first transaction on the transition to substantial usage, and not the potential cost-savings from customers' initial exploratory transactions. Second, to account for the knock-on effect on substantial usage, we need to assume that customers who have not adopted by the last period of our data will do so in the immediately following period. Our results suggest that this transition may often take longer, leading to larger cost savings. Our simulation results show large savings of 12.1% from shortening the delay between login and the first transaction.

Figure 7: Distribution of Cost Savings from Shortening Delays by One Month

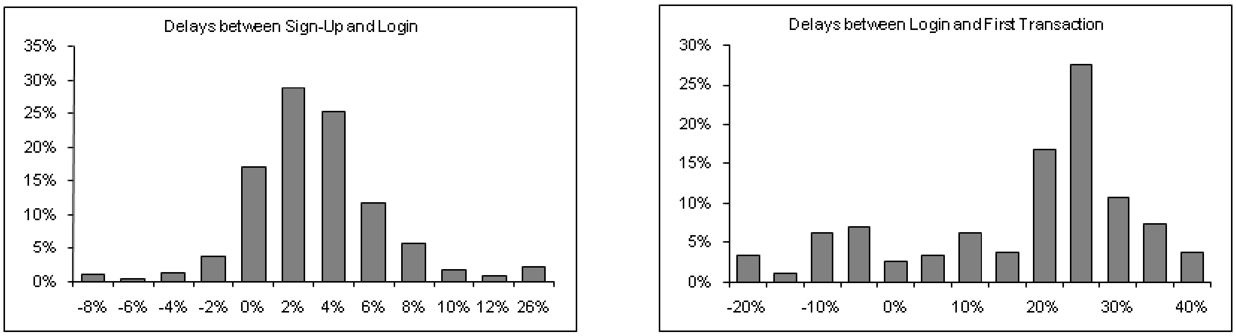


Figure 7 plots the distribution of individual-level cost savings from reducing delays by one month. They show a wider distribution of cost savings at the individual level for speeding up the time between login and the first transaction than for the time between signup and login, where the size of the cost savings is slightly more homogenous. The estimates account only for customers who in our data transitioned to login, respectively a first transaction. Reducing delays will allow more customers to move across the threshold of a login or a first transaction, leading to even larger cost savings for banks.

Our results suggest that moving customers along the adoption funnel speedily can increases ultimate adoption. The staggered adoption process requires that managers pay particular attention to and manage customers' transition along the funnel. Targeting each specific stage in addition to the initial signup has not been a focus in financial institutions to date, but there are relatively easy ways for banks to speed up their customers' movement through the funnel. For example, employees could educate customers on how to use online banking by setting up computers in their branches. Introducing a more differentiated pricing structure could provide financial incentives to customers to move quickly along the adoption funnel. Setting up a webinar with customers who have logged in to demonstrate how to conduct transactions online is an another example of a policy that can ease the transition to the first transaction.

8 Conclusion

Traditionally, researchers have modeled purchases and technology adoption as 'one-off' decisions. This single-stage model works less well for online self-service technologies like banking. In these cases, initial sign-up rarely requires a commitment and rarely has immediate profit implications for firms. Rather, customers have to go through a staggered adoption process, and the firm reaps cost savings only when a customer fully embraces the technology. This paper develops a new empirical methodology to embrace the challenges of estimating a discrete multi-stage adoption process allowing for heterogeneity. The methodology is applicable to a wide variety of services and technologies that involve staggered adoption.

In particular, we explore the empirically observed gap between initial signup and substantial usage. We use the image of an 'adoption funnel' to highlight attrition at four incremental stages: signup, evaluation, trial and substantial usage. We find significant attrition along the adoption funnel: only 24% of customers who sign up for online banking go on to use it substantially. Our results show that exogenous delays in early stages of the adoption funnel explain part of this attrition. These delays significantly reduce a customer's probability of moving to a later stage and to substantial usage. One interpretation, consistent with findings on memory, learning and momentum, is that a delay leads customers to forget what they have already learned about the technology and reduces their motivation to complete the adoption process, which consequently makes them more likely to abandon adoption.

Reducing the time between signup and initial log-in yields significant cost savings. Managers should therefore aim to reduce customer delays in the adoption process by, for example, educating their customers, introducing pricing schemes that provide financial incentives for progress along the funnel, or optimizing the timing of promotions. Our results also suggest that firms could benefit from eliminating stages along the adoption funnel, as Amazon.com did with its 1-click-ordering option to shorten the purchasing process. Similarly, a website increased sales by 40%, simply by allowing customers to bypass the initial registration process (Wroblewski (2008)).

Alternatively, firms can set deadlines by which customers must move to the next stage. Some software companies such as The MathWorks[15] already impose deadlines and require customers to activate the product within a given period after installation. Similarly, E*Trade Financial requires customers to access their account within the first 30 days of sign-up. While customers sometimes use deadlines as commitment devices, externally imposed deadlines are usually more effective than self-imposed deadlines (Ariely and Wertenbroch (2002)). Similar to redemption of coupons with expiration dates (Inman and McAlister (1994)), we would expect that such deadlines lead to a peak in adoption shortly before expiration.

Our results apply to many customer self-service technologies. For example, consumers go through similar stages when they sign up, evaluate, try and subsequently transition (or not) to substantial usage of online grocers. Similarly, consumers may sign up for online account management of reward travel, but never stop receiving paper statements or processing bookings offline. There is also evidence of lack of substantial usage in online entertainment activities such as 'Second Life', where as of July 2007 only 18% of users had logged into the website in the previous 60 days.[16] We believe that in such settings, firms can benefit from directing more attention to collecting data on and actively managing the individual-level adoption funnel.

Our results also highlight that the extensive regulatory efforts to ensure data security could potentially have unintended consequences for the diffusion of online services. In the data we study, the bank initiated the multiple-stage sign-up process in response to a EU directive that required individual authorization for the release of personal data. The impact of such security concerns are not just limited to Europe or to the banking sector. Increasingly complicated processes to protect the security of customers have been suggested in multiple industries such as as healthcare[17] and public services.[18]

[15]The MathWorks, 'Matlab & Simulink Student Version: Frequently Asked Questions', `http://www.mathworks.com/academia/student_version/faq/` (accessed October 9, 2007).

[16]Rose, Frank, 'How Madison Avenue Is Wasting Millions on a Deserted Second Life,' *Wired Magazine 15.08*, July 24, 2007.

[17]See for example, 'Idaho Code 39-1394' which requires written authentication for electronic physician order systems.

[18]See the General Services Administration's Federal Acquisition Service's E-Authentication Services Initiative.

Appendix

A.1 Instrumental Variables

In this section, we estimate a similar specification to our main results in Table 4 by using a two-stage least squares approach. Though this alternative approach does not explicitly account for unobserved heterogeneity in the values of the coefficients, it does allow us to run the standard tests of instrumental variables in a familiar framework. In this estimation approach, we use exogenous variation in school vacations across states and months as instruments to identify the effect of delays. We discuss the background to this strategy and present robustness checks in greater depth.

A.2 Background to Instruments Used

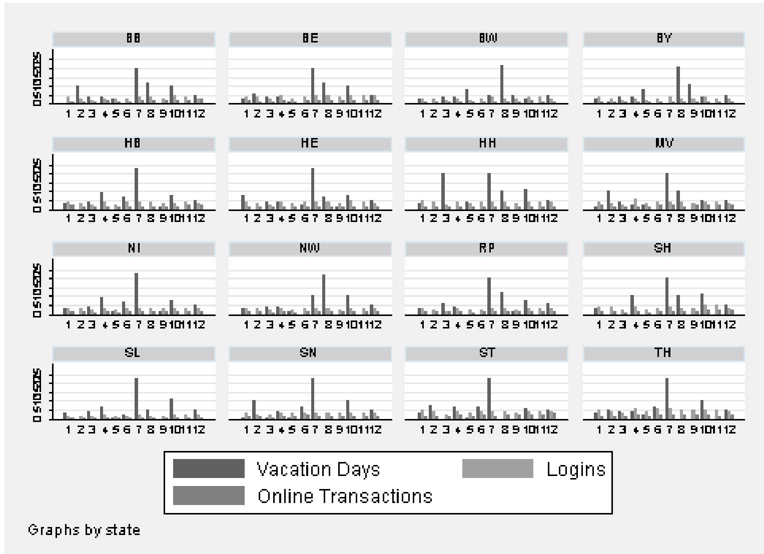
To deal with the challenge of identifying the effect of delays on the adoption process as opposed to other confounding factors, we need a variable that can serve as an exogenous proxy for more general sources of delays. We find such a variable in the number of school vacation days and public holidays across months and across German states.

While educational policy is a responsibility of the individual German states, the federal government coordinates a system of staggered school vacations across states. This government policy aims to reduce traffic congestion during the summer months by ensuring that not everyone goes on vacation at the same time.

This system of staggered school vacations creates an exogenous shifter to a customer's ability to use online banking, because during these state-specified times, many Germans leave their homes to travel. Even with the staggered vacation system, this results in heavy highway congestion at the start and end of the different states' vacations.[19] The automobile club ADAC,

[19] Stroisch, Jörg, "Stau im Urlaub vermeiden", `http://www.zeit.de/reisen/service/stau` (accessed 08/16/07). The fact that such congestion happens despite additional limitation of truck usage of the highway system during the summer months further underlines that many Germans leave their home during vacation periods(Bundesministerium für Verkehr, Bau und Stadtentwicklung, "Lkw-Fahrverbot in der Ferienreisezeit",

Figure A-1: Variation in Vacation over 2002 by State



for example, reports medium to heavy congestion on 95% of the beginning and ending weekends of states' school vacations.[20] The number of days per year spent on vacation is significant. Germans, on average, take 26 vacation days a year (out of the average 5-6 week allowance), compared to the 11 days taken by Americans,[21] and go on roughly 1.6 vacation trips per year.[22] This vacation is often booked up to one year in advance. Figure A-1 displays the extensive variation in vacations both across states and within states over time.

Vacation days affect a customer's ability to engage in the adoption process, because they usually separate her from her computer. This exogenous variation in computer usage has also been exploited by Oberholzer-Gee and Strumpf (2007). They show that German school vacation affect downloads and uploads to P2P sites and use German vacations to instrument for US record sales. In the context of online banking, a customer who leaves her home for vacation is unlikely either to receive the mailing of her online banking details, such as pre-assigned TANs, or to carry

`http://www.bmvbs.de/-,302.2221/Lkw-Fahrverbot-in-der-Ferienre.htm` (accessed 08/16/07).

[20]See `www.adac.de/Verkehr/Staukalender/default.asp` (accessed 08/16/07).

[21]Expedia.com, "2007 International Vacation Deprivation Survey Results", `http://www.expedia.com/vacationdeprivation` (accessed 08/16/07).

[22]Axel Springer Marketing Anzeigen, "Tourismus 2002", `http://www.mediapilot.de/cda` (accessed 08/16/07).

these with her. Government agencies also warn customers against using online banking on public computers at their vacation destination for security reasons.[23] Since recurring transactions, such as utility payments, are typically processed via direct debit, there is little need for online banking to keep up with regular bill payments on vacation.

The decreased accessibility of online banking while traveling implies that a customer who signs up for the service in a vacation-heavy month likely spends longer between signup and login than a customer who signs up at another time, for reasons unrelated to unobservable customer tastes for technology. Time off may similarly affect movement to substantial usage. Similar customers will spend different times between signup and login, and from login to their first transaction, merely because they sign up, or log in, in different months. We use this fact to identify the causal effect of spending longer in the initial stages on the time spent in later stages of the adoption funnel.

A.3 Initial Checks

For vacation and public holidays to be valid instruments, the presence of vacations and public holidays have to be independent of other factors that influence a customer's adoption behavior. For example, it would be problematic if the bank changed its marketing campaigns to match vacations or state public holidays. Conversations with the bank assured us, however, that the bank conducted such marketing efforts at a national level only. We also do not observe clustering of customer signups for online banking just prior to or immediately after school vacations. Across customers, 22.0% sign up for the service during vacations, 13.9% within 10 days of the start of the closest vacation and 14.1% within 10 days of the end of the closest vacation. The remaining 50% sign up outside of vacation periods.[24] This suggests that customers do not take into account planned vacations in deciding when to sign up for online banking. Table A-1 presents further

[23] Bundesamt für Sicherheit in der Informationstechnik, "Brennpunkt: IT-Betrueger machen keinen Urlaub", `http://www.bsi-fuer-buerger.de/brennpunkt/urlaub.htm` (accessed 08/16/07).

[24] Across states, schools close for vacation 20% of the time. 34% of vacation days fall within 10 days of a school vacation period.

Table A-1: Offline Transaction Activity before and after vacation-heavy months

# Days Off Previous Month	Observations	Mean Transactions	Standard Error
<5 Days Off	191	8.608	0.144
≥5 Days Off	193	8.689	0.180
Difference		-0.081	0.231
P-value that difference is not significantly different from zero: 0.726			

# Days Off Subsequent Month	Observations	Mean Transactions	Standard Error
<5 Days Off	191	8.229	0.135
≥5 Days Off	193	8.309	0.171
Difference		-0.079	0.219
P-value that difference is not significantly different from zero: 0.718			

Table A-2: Proximity of signup decision to vacations

	N	Age ≤ 40	Age > 40	Men	Women
Vacation close to signup date					
(1) Signup in vacation	756	67.6	32.4	52.4	43.7
(2) Signup < 10 days before vacation	479	69.5	30.5	50.7	43.4
(3) Signup < 10 days after vacation	487	70.8	29.2	51.1	45.0
(4) Signup at other times	1721	66.7	33.3	51.3	43.1
Difference bw. (1) and (4) significant		no	no	no	no
Difference bw. (2) and (4) significant		no	no	no	no
Difference bw. (3) and (4) significant		no	no	no	no
Days of vacation in signup month					
(5) Vacation below median (≤ 3 days)	1847	67.2	32.8	49.9	44.1
(6) Vacation above median (> 3 days)	1595	68.7	31.3	53.2	43
Difference bw. (5) and (6) significant		no	no	no	no

* p<0.10, ** p<0.05, *** p<0.01.

evidence that suggests that offline banking activity is not correlated with vacation days.

We also checked whether signup before, during or after vacations is correlated with customer characteristics. Our assumption that our instrument is a random treatment across customers could be violated if customers who likely have higher internet usage rates, such as younger people, were concentrated in states that have a higher number of vacations. Table A-2 compares the shares of customers who sign up close to or during a vacation or who have an above-median amount of vacation in their signup month. Signup behavior does not differ significantly in relation to vacations for different demographic groups. This suggests that vacation in the signup month is not correlated with observable customer characteristics.

This evidence suggests that our vacation days instrument is uncorrelated with unobservables

Table A-3: Determinants of Time between Signup and Login

	Coefficient	**Robust Std. Error**	**Sig.**
Vacation Days, Signup Month	3.3616	0.3821	***
Public Holidays, Signup Month	-21.5539	2.7267	***
Branches in Zip	33.6380	4.8357	***
Month-Year Dummies	Yes		
State Dummies	Yes		
Observations	5691		

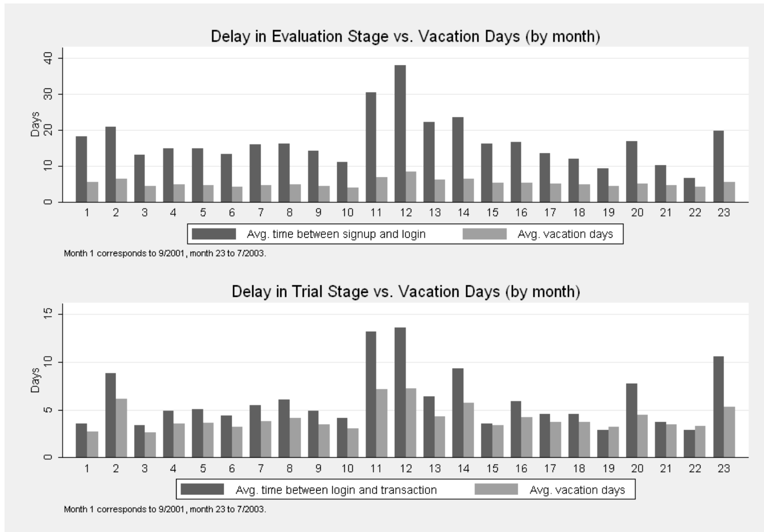
Dependent Variable: Days between signup and first login. Sample: Customers who have logged in but not yet performed an online transaction. * p<0.10, ** p<0.05, *** p<0.01.

that may influence adoption behavior. There is also similar evidence that public holidays are uncorrelated with the unobservables that may influence adoption behavior. In addition, vacations are strongly correlated with the endogenous variable: Table A-3 displays evidence for the strong positive relationship between the time a customer spends between signup and login and vacation days, and for a negative relationship between this length of time and public holidays. Both the number of vacation days and of public holidays in the signup month have statistically significant effects on the average time spent in moving from signup to login. We find that additional days of vacation increase the time delay, consistent with customers being away from their home computers. We find that additional public holidays decrease the time between signup and login consistent with the thesis that people use these days to catch up with chores. Increasing the number of vacation days in the customer's signup month by one standard deviation of 6.21 days increases the average delay between signup and login by 20.87 days.

We also depict graphically the relationship between the residual time spent in the first two stages of the adoption funnel, after removing the effect of the other explanatory variables in Table A-3, and the number of vacation days in the signup or login month. Figure A-2 illustrates the co-movement in delays and vacation times, suggesting that vacation time at the time of earlier decisions indeed shifts the time spent moving to a later stage of the adoption process.

We find that a customer's probability of moving to the next stage is affected by the number

Figure A-2: Relationship between residual time between signup and login or between login and online transaction and vacation days



of vacation days at the time she completed her previous transition. Vacation days in the signup month decrease a customer's probability of logging in for the first time (-0.009, p-value <0.001 for vacation), controlling for branch density, and state and month dummies. We find similar results for a customer's probability of moving from her first login to her first online transaction.

A.4 Full Instrumental Variables Model

Table A-4 presents our full instrumental variable results. Because a traditional logit does not lend itself easily to an instrumental variable estimation, we use a probit specification. Though this specification does not explicitly account for heterogeneity in the values of the coefficients as is the case in Table 4, it does allow us to run the standard tests of instrumental variables in a familiar framework.

In this specification the dependent variable covers both whether the customer progresses to stage 2, that is, she makes a first online transaction, and whether the customer progresses to stage 3, that is, she conducts most of her transactions online. The delay between signup and login and its interaction terms are instrumented by both the number of vacations and public holidays in the month that the customer signed up. The delay between login and first online transaction and its interaction terms are instrumented by the number of vacations and public holidays in the month that the customer first made a login. To control for heterogeneity at the individual level, we clustered our standard errors at the group level across both stages in the funnel using a bootstrap. The results are similar to our main specification. Delays in both stages in the adoption process slow down progress to adoption.

Since we have six endogenous variables and do not explicitly incorporate individual level heterogeneity in the coefficients, our estimates are less precise than in our main specification in Table 4. To check the validity of our estimation approach, we conduct familiar tests of instrument validity. The Anderson-Rubin Wald F-test for the first stage of estimation showed that the instruments in all cases for the transition to first transaction and to substantial usage were jointly significant at the 1% level for the first stage. The insignificance of the Sargan test for over-identification suggests that the specification is not over-identified.

A.5 Further Robustness Checks

In the main analysis, we define arrival at substantial usage as the point at which a customer switches 50% of her transaction activity to online banking. To check the robustness of our results, we estimated our model with an alternative less restrictive definition of substantial usage where we define arrival at substantial usage as shifting 40% of transactions online. The estimation results in Table A-5 confirm the significantly negative effect of delays in earlier stages on later-stage adoption. We find a similar degree of heterogeneity as in our main specification, with a probability for the negative class 1 of 69.9% (versus 77.8% in the main specification).

Table A-4: Instrumental Variables Estimates for Integrated Model

	Marg. Effect	**Std. Error**	**Sig.**
Stage 2: Probability to Transition from Login to First Transaction			
Delay between Signup and Login	-0.493	0.221	**
Delay between Signup and Login*Male	0.025	0.012	**
Delay between Signup and Login*Age	-0.004	0.002	**
Age	0.003	0.002	
Male	-1.001	0.471	**
Brokerage	0.062	0.058	
Public Holidays	0.038	0.028	
Vacation Days	-0.001	0.002	
Stage 3: Probability to Transition from First Transaction to Substantial Usage			
Delay between Login and First Transaction	-0.017	0.098	
Delay between Login and First Transaction*Male	-0.002	0.006	
Delay between Login and First Transaction*Age	0.001	0.001	
Age	-3.2E-04	0.001	
Male	0.056	0.103	
Brokerage	-0.016	0.017	
Public Holidays	-0.009	0.014	
Vacation Days	-0.001	0.001	
Stage Specific Monthly Fixed Effects	Yes		
Stage Specific State Fixed Effects	Yes		
Observations	22643		
Log-Likelihood	-13694.8		
BIC	28312.2		
AIC	27573.7		
Sargan Test Stat	0.187		
Sargan Test P-Value	0.666		
AR Wald Statistic	35.62		
AR Wald P-Value	9.2E-06		

Instrumental variables used for all delay variables. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.
Bootstrapped standard errors, clustered at the group level.

Table A-5: Main Results for Multi-Stage Adoption Process: Alternative Definition of First Arrival to Substantial Usage

Variable	**Coefficient**	**Std. Error**	**Sig.**	**Marg. Eff.**
Stage 1: Probability to Transition from Sign-Up to Login				
Intercept stage 1 (class 1)	0.178	0.144		0.043
Intercept stage 1 (class 2)	0.022	0.176		0.005
Male	0.066	0.083		0.016
Age	0.000	0.003		-2.4E-05
Brokerage account	0.110	0.087		0.026
Vacation Days in Month	-0.016	0.007	**	-0.004
Public Holidays in Month	-0.002	0.041		-0.001
Time spent in stage 1	-0.126	0.016	***	-0.030
Stage 2: Probability to Transition from Login to First Transaction				
Delay between Signup and Login (class 1)	-0.413	0.104	***	-0.059
Delay between Signup and Login (class 2)	-0.168	0.100	*	-0.024
(Delay between Signup and Login) * Male	0.142	0.059	**	0.020
(Delay between Signup and Login) * Age	0.006	0.002	**	0.001
Intercept stage 2 (class 1)	2.248	0.278	***	0.324
Intercept stage 2 (class 2)	1.611	0.320	***	0.231
Male	-0.450	0.157	***	-0.065
Age	-0.020	0.006	***	-0.003
Brokerage account	-0.405	0.095	***	-0.058
Vacation Days in Month	-0.004	0.008		-0.001
Public Holidays in Month	-0.022	0.044		-0.003
Time spent in stage 2	-0.555	0.027	***	-0.133
Stage 3: Probability to Transition from First Transaction to Substantial Usage				
Delay between Login and First Transaction (class 1)	-0.444	0.195	**	-0.023
Delay between Login and First Transaction (class 2)	0.100	0.135		0.007
(Delay between Login and First Transaction) * Male	0.245	0.138	*	0.014
(Delay between Login and First Transaction) * Age	0.003	0.004		1.9E-04
Intercept stage 3 (class 1)	-1.346	0.302	***	-0.071
Intercept stage 3 (class 2)	-1.807	0.291	***	-0.126
Male	-0.075	0.195		-0.004
Age	-0.002	0.007		-9.9E-05
Brokerage account	-0.377	0.109	***	-0.022
Vacation Days in Month	0.008	0.008		4.5E-04
Public Holidays in Month	-0.025	0.047		-0.001
Time spent in stage 3	-0.166	0.013	***	-0.040
Stage Independent				
Parameter for class probability, θ_c	0.841			
Implied class probablility (class 1)	0.699			
Log-Likelihood	-5451.9			
Observations	1297			
McFadden R-Squared	0.271			
Adjusted McFadden R-Squared	0.266			

Arrival at substantial usage defined as the first month in which a customer conducts 40% or more transactions online. * p<0.10, ** p<0.05, *** p<0.01.

References

Allison, P. (1982). Discrete-time methods for the analysis of event histories. *Sociological Methodology 13*, 61–98.

Allison, P. D. and N. A. Christakis (2006). Fixed effects methods for the analysis of nonrepeated events. *Sociological Methodology 36*(1), 155–172.

Ariely, D. and K. Wertenbroch (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological Science 13*(3), 219–224.

Bahrick, H. (1979). Maintenance of knowledge: Questions about memory we forgot to ask. *Journal of Experimental Psychology: General 108*(3), 296–308.

Bass, F. (1969). A new product growth model for consumer durables. *Management Science 15*, 215–227.

Beal, G., E. Rogers, and J. Bohlen (1957). Validity of the concept of stages in the adoption process. *Rural Sociology*.

Davis, F. (1989). Perceived usefulness, perceived ease of use, and user accceptance of information technology. *MIS Quarterly 13*, 319–340.

Dhar, R., J. Huber, and U. Khan (2007). The shopping momentum effect. *Journal of Marketing Research 44*(3), 370–378.

Goettler, R. and K. Clay (2006). Tariff Choice with Consumer Learning: Sorting-Induced Biases and Illusive Surplus. Mimeo, Carnegie-Mellon University.

Goldfarb, A. and J. Prince (2007). Internet Adoption and Usage Patterns are Different: Implications for the Digital Divide. Mimeo, University of Toronto.

Gordon, R. (2000). Does the "new economy" measure up to the great inventions of the past? *The Journal of Economic Perspectives 14*(4), 49–74.

Gupta, S. (1991). Stochastic models of interpurchase time with time-dependent covariates. *Journal of Marketing Research 28*, 1–15.

Heckman, J. and B. Singer (1984, March). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica 52*(2), 271–320.

Hitt, L. and F. Frei (2002). Do Better Customers Utilize Electronic Distribution Channels? The Case of PC Banking. *Management Science 48*(6), 732–748.

Inman, J. and L. McAlister (1994). Do coupon expiration dates affect consumer behavior? *Journal of Marketing Research 31*, 423–428.

Jenkins, S. P. (1995, February). Easy estimation methods for discrete-time duration models. *Oxford Bulletin of Economics and Statistics 57*(1), 129–38.

Kalish, S. (1985). A new product adoption model with price, advertising, and uncertainty. *Management Science 31*(12), 1569–1585.

Kamakura, W. A. and G. J. Russell (1989, Nov). A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research 26*(4), 379–390.

Lambrecht, J. (2005). *Einsatz des Online-Bankings zur Steigerung des Kundenwerts.* Ph. D. thesis, Shaker Verlag, Aachen.

Meuter, M., M.-J. Bitner, A. Ostrom, and S. Brown (2005). Choosing among alternative service delivery modes: An investigation of customer trial of self-service technologies. *Journal of Marketing 69*(2), 61–83.

Meyer, B. D. (1990, July). Unemployment insurance and unemployment spells. *Econometrica 58*(4), 757–82.

Oberholzer-Gee, F. and K. Strumpf (2007). The effect of file sharing on record sales: An empirical analysis. *Journal of Political Economy 115*(1), 1–42.

Rogers, E. (2003). *Diffusion of Innovations* (Fifth ed.). Free Press.

Seetharaman, P. B. and P. K. Chintagunta (2003, July). The proportional hazard model for purchase timing: A comparison of alternative specifications. *Journal of Business & Economic Statistics 21*(3), 368–82.

Soman, D. (2003). Prospective and retrospective evaluations of experiences: How you evaluate an experience depends on when you evaluate it. *Journal of Behavioral Decision Making 16*, 35–52.

Speier, C., J. Valacich, and I. Vessey (1999). The effects of interruptions, task complexity, and information presentation on computer-supported decision-making performance. *Decision Sciences 30*(2), 771–779.

Speier, C., J. Valacich, and I. Vessey (2003). The influence of task interruption on individual decision making: An information overload perspective. *Decision Sciences 34*(4), 337–360.

Tellis, G., S. Stremersch, and E. Yin (2003). The international takeoff of new products: The role of economics, culture and country innovativeness. *Marketing Science 22*(2), 188–208.

Van Den Bulte, C. and G. Lilien (2007). A two-stage model of innovation adoption with partial observability: Model development and application. Mimeo, University of Pennsylvania.

Van Den Bulte, C. and S. Stremersch (2004). Social contagion and income heterogeneity in new product diffusion: A meta-analytic test. *Marketing Science 23*(4), 530–544.

Wroblewski, L. (2008). *Web Form Design.* Rosenfeld Media.

Wuebker, G. and P. Hardock (2002). Online banking: Weit verbreitet, doch kaum genutzt? *Die Bank 6.*