ESSAYS ON THE ECONOMICS OF BROADBAND NETWORKS

by

JACOB BRADLEY MALONE

(Under the Direction of Scott Atkinson)

ABSTRACT

The Internet has become an integral part of everyday life. Now, more than ever, it is important to understand these markets to best guide and inform public policy. This dissertation uses proprietary, high-frequency, subscriber-level data from various North American Internet Service Providers to study the demand side of residential broadband markets. In Chapter 1, I study how subscribers are currently using the Internet with a particular focus on the current role and importance of online video. In Chapter 2, I explore the implications of usage-based pricing and its ability to improve efficiency in broadband markets. In Chapter 3, I estimate subscriber demand of residential broadband where variation in network congestion and prices are included.

INDEX WORDS: Broadband, Internet, Demand
Essays on the Economics of Broadband Networks

by

Jacob Bradley Malone

B.A., University of Georgia, 2011
M.A., University of Georgia, 2011

A Dissertation Submitted to the Graduate Faculty
of The University of Georgia in Partial Fulfillment
of the
Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2016
ESSAYS ON THE ECONOMICS OF BROADBAND NETWORKS

by

JACOB BRADLEY MALONE

Approved:

Major Professor: Scott Atkinson

Committee: Jonathan Williams
Meghan Skira
Thomas Quan

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2016
Acknowledgments

This dissertation is the product of numerous inputs over the years, which there are too many to name, but for each I am immensely appreciative and grateful.

First, I want to thank my wife Carol for her never-ending support and encouragement. Without her, this dissertation is not possible.

Next, I want to thank Professor Jonathan Williams for his commitment, patience, and guidance in both my undergraduate and graduate careers. His advisement has made me the careful and inquisitive researcher I am today, and the quality of his research is what I aspire mine to be. I also thank Professors Atkinson, Skira, and Quan for supporting my research, providing comments, and serving on my dissertation committee.

I also want to thank Ron Reuss and Sanjay Patel for sharing their expertise on broadband markets with me, and the Internet Service Providers who graciously made available the data foundational to this research.

Finally, I thank my family, particularly my parents, for making so many opportunities available to me and instilling in me the desire and skills needed to complete this dissertation.
Contents

Acknowledgments iv

1 A Snapshot of the Current State of Residential Broadband Networks 1
   1.1 Introduction .............................................. 1
   1.2 Data ...................................................... 8
   1.3 Overview of Online Activity ............................. 10
   1.4 Cord Cutting and Pay TV Substitution ................... 19
   1.5 Conclusion .............................................. 23

2 Do Three-Part Tariffs Improve Efficiency in Residential Broadband Networks? 27
   2.1 Introduction .............................................. 27
   2.2 Data ...................................................... 32
   2.3 Monthly Usage .......................................... 36
   2.4 Intra-Day and Intra-Month Usage ....................... 40
   2.5 Conclusion .............................................. 46

3 The Tragedy of the Last Mile: Congestion Externalities in Broadband Networks 51
   3.1 Introduction .............................................. 51
# List of Figures

1.1 Monthly Usage by Quantile and Group ........................................ 14
1.2 Average Hourly Usage .............................................................. 15
1.3 Transition Matrices of Group Usage ............................................ 26

2.1 CMTS Relationship to Subscriber .............................................. 33
2.2 Yearly Growth and Plan Upgrade Incidence .................................. 36
2.3 Usage Distributions ................................................................. 37
2.4 Usage Transition Surface ......................................................... 38
2.5 Average Intra-Day Subscriber Bandwidth Usage .......................... 41
2.6 Hourly Usage as a Percent of Daily Total ................................... 42

3.1 Internet Plan Features ............................................................... 59
3.2 Average Usage by Hour and Direction ....................................... 61
3.3 Statistics of Usage as a Percentage of Daily Total ....................... 61
3.5 Hourly Aggregate Packet Statistics ............................................ 64
3.6 Average Hourly Subscriber Packet Loss .................................... 65
3.7 Weekly Node Utilization Statistics ............................................. 67
3.8 Heatmap of Peak Packet Loss Transitions .................................. 69
3.9 Estimated Marginal Distributions of Model Parameters ............... 80
3.10 Estimated Joint Distributions of Model Parameters . . . . . . . . . . . . . . 82
# List of Tables

1.1 *DPI Application Groups* .................................................. 9  
1.2 *Descriptive Statistics of Subscribers* .................................. 12  
1.3 *Average Monthly Usage by Group* ......................................... 13  
1.4 *Average Hourly Total Usage During Off-Peak and Peak Hours* ......... 16  
1.5 *Average Hourly Group Usage During Off-Peak and Peak Hours* ........ 17  
1.6 *Normalized Usage for Peak and Off-Peak Hours* ........................... 18  
1.7 *Percentage of Cord Cutters by Decile* .................................... 20  
1.8 *Unconditional Changes in Average Daily Usage by Group* .............. 21  
1.9 *Unconditional Changes in Average Daily Usage by Video Type* ......... 23  
1.10 *Conditional Changes in Average Daily Usage by Video Type* .......... 24  

2.1 *Summary Statistics of Plan Details* ....................................... 34  
2.2 *Summary Statistics of Monthly Usage* .................................... 48  
2.3 *Summary Statistics of Monthly Revenue* .................................. 49  
2.4 *Main Regression Estimates* ................................................. 50  

3.1 *Daily Usage Distributions by Internet Plan Tier* .......................... 60  
3.2 *Changes in Node Utilization and Packet Loss After Node Split* ........ 68  
3.3 *Changes in Daily Usage After Node Split* .................................. 68  
3.4 *Transition Matrix of Peak Packet Loss* ................................... 69
3.5 Descriptive Statistics for Types ........................................ 81
3.6 Counterfactual Results from Eliminating Network Congestion .... 83
Chapter 1

A Snapshot of the Current State of Residential Broadband Networks

1.1 Introduction

The way in which consumers use the Internet is changing rapidly, such that traditional activities like web browsing, email, and file sharing now represent only a modest share of aggregate traffic. The major driver of this trend is the popularity of over-the-top-video (OTTV) services like Netflix, Amazon, Sling TV, and YouTube, which collectively represent the vast majority of Internet traffic and an increasingly better substitute for traditional pay TV services. The dramatic rise of OTTV services has led to major changes in the telecommunications sector and greater attention to several important and ongoing public policy debates like net-neutrality, usage-based pricing, merger review, and municipal broadband.

In this paper, I provide an essential component to inform these policy debates: an in-depth descriptive analysis of the rapidly changing way in which consumers use the Internet from a representative sample of consumers. At the core of my contribution are unique data from the spring and summer of 2015 that I acquired from a North American ISP. These data
include detailed disaggregated high-frequency information on individuals’ Internet usage, as well as information on billing and subscriptions to other services like traditional linear TV and wireline phone service. Using these data, I provide an analysis of current usage patterns with a focus on OTTV services. At the individual consumer level, I provide insight into temporal patterns in usage within the day, measure persistence in the level and composition of usage across days, and contrast usage patterns for consumers that subscribe to traditional pay TV services (i.e., purchase a bundle of services from the operator) to those that do not. My data also permit me to examine how usage changes around the time that a consumer “cuts the cord” and cancels traditional linear TV service, a trend that is rapidly accelerating as OTTV services improve.

I find that OTTV and streaming services are now over 62% of all residential internet traffic, while web browsing is nearly 27%. Differences in usage levels across subscribers are very large, and disproportionately due to OTTV usage. For example, the 25th percentile subscriber averages 14.3 GBs per month, while the 75th percentile subscriber averages 136.4 GBs. The 99th percentile user averages nearly 600 GBs of total usage each month, including over 300 GBs of OTTV services alone, which is the equivalent of four full-length HD movies per day. Interestingly, I observe that almost every household has some engagement with OTTV, which suggests that the market penetration of these services is increasing rapidly and now represents a serious competitor with traditional linear TV services. Despite findings in other related industries that suggest subscribers’ usage is unpredictable and erratic (e.g., cellular phone usage), my data reveals very strong persistence in usage habits. Specifically, I observe that past behavior at the daily (monthly) level is almost a perfect predictor of behavior in subsequent days (months), for both overall usage and individual applications. Finally, I find huge disparities in usage between those subscribers with traditional TV service, and those without. Bundled subscribers average 81.1 GBs per month, while non-bundled subscribers average nearly twice as much, 159.2 GBs. Nearly all of this difference is attributable
to OTTV services. The high frequency of my data permits me to examine the behavioral change of those subscribers immediately before and after dropping TV service. The change in average daily usage is about 34%, or 1.3 GBs, despite substantial engagement of OTTV services prior to dropping traditional linear TV service. For these “cord cutters”, there is a large and immediate increase in the use of OTTV services of which 0.7 GBs is Netflix alone. These descriptive insights are important inputs into a number of ongoing public policy debates.

The net-neutrality debate over the equal treatment of content on the Internet has garnered national attention recently. The rise of OTTV services put pressure on regulators (i.e., the FCC and DOJ) to more closely monitor how ISPs treat the traffic generated by these services. Specifically, when a broadband customer uses third-party OTTV services, ISPs receive no additional revenue, only greater network costs, and also often lose a traditional linear TV subscriber. Thus, there is incentive to differentially treat these sources of traffic to lessen their attractiveness. This tension, along with allegations by OTTV service providers\footnote{See WSJ (2014) for allegations by Netflix of ISPs degrading interconnections to slow their traffic during negotiations, and Clark (2014) for research that shows Netflix selectively chose routings for its traffic that degraded performance of their services during negotiations.}, led, in part, to the FCC’s Open Internet rules being adopted on February 26, 2015 and implemented on June 12, 2015.

The Bright Line Rules from this order include\footnote{See https://www.fcc.gov/openinternet}:

- No Blocking: broadband providers may not block access to legal content, applications, services, or non-harmful devices.
- No Throttling: broadband providers may not impair or degrade lawful Internet traffic on the basis of content, applications, services, or non-harmful devices.
- No Paid Prioritization: broadband providers may not favor some lawful Internet traffic over other lawful traffic in exchange for consideration of any kind, in other words, no
“fast lanes.” This rule also bans ISPs from prioritizing content and services of their affiliates.

This order, whose legal authority is rooted in Title II of the Communications Act and Section 706 of the Telecommunications Act of 1996, represents the strongest form of net neutrality.3 It prevents ISPs from numerous practices that may have welfare enhancing properties like active network management practices that prioritize congestion-sensitive applications (e.g. OTTV). It could also limit future agreements like the landmark one reached between Comcast and Netflix in February 2014 where Netflix agreed to pay Comcast for faster and more reliable access to Comcast’s subscribers (NYT 2014). The agreement essentially provides Netflix with its own pipe or “on ramp” to Comcast subscribers, rather than “paid prioritization” where Netflix traffic would receive preferential treatment among other applications in the same pipe. The latter being regarded by most as a clear violation of net-neutrality principles. There is an extensive theoretical literature on these topics (e.g. Economides (2015), Economides and Tag (2012b), Economides and Tag (2012a), Economides and Hermalin (2012), Choi et al. (2013b), Choi et al. (2013a)), which are helpful for understanding the implications of this regulatory change, but very few provide unambiguous predictions regarding welfare. I view my empirical analysis that identifies trends in consumers’ usage patterns and preferences as a first step towards providing the necessary information to make firmer predictions of the impact of the new regulatory environment, including how both ISPs and content providers will respond.

One strategy that ISPs are increasingly adopting in response to the popularity of OTTV services, and the new regulatory environment, is usage-based pricing (UBP) for broadband services.4 The most common form of UBP is a simple 3-part tariff, where subscribers pay

---

3This legal footing for the FCC’s order is currently being challenged on a number of fronts.
4Economides (2015) show that ISPs also have a strategic incentive to implement UBP to spur upstream competition among content providers that generates a surplus to consumers, which can then be extracted through higher fees.
a fixed fee for a monthly usage allowance, and pay a marginal price for usage in excess of
the allowance. Malone et al. (2014) and Nevo et al. (2016) study the impact of UBP and
find mixed results: UBP is effective at removing low-value usage from broadband networks
at the cost of a small transfer from consumers to the ISP. However, the rudimentary form
of UBP often implemented by most ISPs removes traffic at all times of the day, including
off-peak hours when usage is essentially costless to the ISP. This results in a pure welfare loss
since the surplus from the costless usage is neither enjoyed by the consumer nor extracted as
revenue by the firm. This suggests that even very mild forms of peak-use pricing (e.g. only
counting usage during peak hours against the consumer’s usage allowance) may be welfare
enhancing.

Implementation of the most simplistic forms of peak-use pricing by ISPs would likely
be successful at spurring changes by content providers that would yield substantial welfare
gains. I show that more than 70% of traffic during peak hours, when traffic is over five
times higher than the trough in the early morning hours, is OTTV and streaming services.
Currently, there is almost no incentive for OTTV service providers to locally cache con-
tent for their subscribers, other than to facilitate instant access to content. This was the
motivation behind Amazon’s Advanced Streaming and Prediction (ASAP) feature, which
predictably and continually caches content that the consumer is likely to watch on the Ama-
zon Fire device. As Amazon learns the usage patterns and preferences of the consumer, the
ASAP feature becomes more effective and reliable at queuing up the correct content. Thus,
the Amazon ASAP feature functions much like caching done by Akamai Content Delivery
Networks, but just at a more local level: the user’s Amazon Fire device.

Amazon’s ASAP feature was designed to improve consumers’ experience, not to lower
network costs (some of which would be passed on to consumers through lower prices), and
as a result I find that the temporal patterns in usage across the day for Amazon are nearly
identical to all other OTTV services. If ISPs were to implement a simplistic form of peak-use
pricing rather than the three-part tariffs currently being used, one would expect these types of technology to be highly effective at caching the correct content locally during off-peak hours. For example, Netflix, a big opponent of any form of usage-based pricing including peak-use pricing since its traffic represents about one-third of all traffic, has suggested that it can almost perfectly predict whether a customer will watch the next show in a series (Chicago Tribune 2015). If this is the case and ISPs take the sensible step of implementing peak-use pricing, I should see many more OTTV services using predictive caching during off-peak hours, or even letting subscribers to their services selectively cache programming that they intend to watch soon. This would be very similar in function to DVR services currently offered by DirecTV, for example, and would enhance the welfare of all parties through improved quality of OTTV services and lower costs to the ISP.

My analysis is also informative for the ongoing merger reviews resulting from the rapid consolidation in the telecommunications industry. The first of which was the proposal by Comcast to purchase Time Warner Cable, that ultimately was called off due to resistance by the DOJ and FCC. A merger of ATT and DirecTV was finally approved with some conditions, which actually included strict limits on the use of usage-based plans. There are also now pending mergers between Charter Communications and Time Warner, and Altice and Cablevision. Interestingly, some of the debate surrounding these mergers involved the intensity of competition for traditional linear TV services. My results show that competition from OTTV services, in addition to satellite television, is quite intense. Nearly every household in my data with an Internet connection uses OTTV services in some capacity, and the rate of cord cutting is accelerating with major network cost implications for ISPs. I expect this to exert pressure on margins for pay TV and result in substantial changes to bundling practices, like the introduction of “skinny bundles” much like Sling TV currently offers. In addition to merger review, my analysis is useful for the ongoing debate around municipal broadband networks. As video service transitions away from multicasting towards OTTV
offerings, it will be important that those in rural areas have reasonable options for consuming video. Satellite TV is certainly one option that is often available, but having access to a high-speed connection will be crucial. Thus, municipal broadband networks, or incentives for private ISPs to expand the reach of their networks, will be an important component to ensuring equal access to content for consumers.

My work complements and contributes to a number of important literatures on the economics of the Internet. Most directly, my work builds on numerous previous studies of demand for residential broadband. Varian (2002) and Edell and Varaiya (2002) run experiments, where users face varying prices for different allowances and speeds. Goolsbee and Klenow (2006) use data on individuals’ time spent on the Internet and earnings to estimate consumer benefit from residential broadband, assuming an hour spent on the Internet is an hour of forgone wages. Lambrecht et al. (2007) use monthly consumption data from a German ISP to study the role of uncertainty in consumers’ selection of usage-based plans. Several additional papers (Dutz et al. (2009); Rosston et al. (2010); Greenstein and McDevitt (2011)) estimate the economic value of broadband internet using plan choice data. Additionally, my work provides further insight into the likely impact of different forms of usage-based pricing. Much of the debate on the usage and welfare implications of usage-based pricing has been theoretical (Mackie-Mason and Varian (1995); Bauer and Wildman (2012); Odlyzko et al. (2012); Economides and Hermalin (2015)), and has not been informed by data. My analysis builds on Nevo et al. (2016) and Malone et al. (2014)) to get closer to answering many of the questions raised by FCC’s Open Internet Advisory Committee regarding the impact of usage-based pricing (OIAC (2013)). In summary, the biggest difference between my work and previous related studies is the detail of the data, which permits for a much richer analysis of Internet usage. This detail is crucial given the nature of current policy debates.
The remainder of the paper is as follows. In Section 1.2, I discuss the source and representativeness of my data, and the process by which I merged the different sources, before providing basic summary statistics. I present the results of my analysis in Sections 1.3 and 1.4, and conclude in Section 1.5.

1.2 Data

The data for my analysis comes from a representative sample of broadband subscribers. The data includes detailed information from a large market of a North American ISP for five months, April 2015 to August 2015. The average income in this market is within 5% of the national average, and the demographic composition is nearly identical to the overall US population. Thus, I expect the insights from my analysis to have external validity to other similarly representative markets, and the overall US population.

I have three primary sources of data. The first are Internet Protocol Detail Records (IPDR) data which includes a variety of information: modem MAC address (hashed to preserve anonymity), bytes and packets passed in both the downstream and upstream direction, packets dropped or delayed, and information on the topology of the network (i.e., how subscribers are connected to one another). IPDR data is typically collected at the 15-minute frequency, but my data has been aggregated to an hourly frequency. The second source of data is deep packet inspection (DPI) from one of the two major vendors (i.e., Procera and Sandvine). The DPI data provides the byte counts by MAC address (hashed to preserve anonymity) for each application (e.g., Netflix) or protocol (e.g., File Transfer Protocol) at a five-minute maximum frequency. DPI data provided to me was aggregated to the hourly level. The final source of data for my analysis is billing tables. The billing tables include: the subscriber’s monthly bill, details of the services subscribed to (e.g., characteristics of broadband plan like speed in Mb/s along with video and phone service indicators), a (hashed)
Table 1.1: DPI Application Groups

<table>
<thead>
<tr>
<th>Groups</th>
<th>Description (Examples)</th>
<th>% of All Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administration</td>
<td>System administrative tasks (STUN, ICMP)</td>
<td>1.19</td>
</tr>
<tr>
<td>Backup</td>
<td>Online storage (Dropbox, SkyDrive)</td>
<td>0.58</td>
</tr>
<tr>
<td>Browsing</td>
<td>General web browsing (HTTP, Facebook)</td>
<td>26.70</td>
</tr>
<tr>
<td>CDN</td>
<td>Content delivery networks (Akamai, Level3)</td>
<td>2.95</td>
</tr>
<tr>
<td>Gaming</td>
<td>Online gaming (Xbox Live, Clash of Clans)</td>
<td>3.06</td>
</tr>
<tr>
<td>Music</td>
<td>Streaming music services (Spotify, Pandora)</td>
<td>3.40</td>
</tr>
<tr>
<td>Sharing</td>
<td>File sharing protocols (BitTorrent, FTP)</td>
<td>0.20</td>
</tr>
<tr>
<td>Streaming</td>
<td>Generic media streams (RTMP, Plex)</td>
<td>6.26</td>
</tr>
<tr>
<td>Tunnelling</td>
<td>Security and remote access (SSH, ESP)</td>
<td>0.07</td>
</tr>
<tr>
<td>Video</td>
<td>Video streaming services (Netflix, YouTube)</td>
<td>55.47</td>
</tr>
<tr>
<td>Other</td>
<td>Anything not included in above groups</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Note: This table lists groups used to summarize individual applications found in the DPI data. Each group characterizes a different type of online activity. Applications are placed into a group if they are within the top 175 applications by either total usage or frequency. Otherwise, an application is labeled as Other. Streaming is separated from Music and Video because these applications cannot be identified as audio or video streams, whereas applications like Spotify or Netflix can. The final column reports the percentage of all subscriber usage generated by each group and sums to 100%.

customer key that can be linked to a (hashed) MAC address to identify a household even when hardware changes, like a modem upgrade. The common keys between the different data sets, hashed customer keys and MAC addresses, permits me to link the data sources for my analysis.

Even though IPDR and DPI data both cover subscriber usage, they differ in scope. IPDR only reports how many bytes a subscriber sends and receives within an hour, while the DPI data record hourly usage by application (e.g., Netflix) or protocol (e.g., File Transfer Protocol (FTP)). The DPI data do not record the direction of the traffic, while IPDR does. By linking the two sources, however, I can gain a complete picture of usage.

Both the IPDR and DPI data contain byte counts for each subscriber at an hourly frequency, but the aggregation process results in small discrepancies between the two sources.
To resolve the differences, I treat IPDR as the authoritative source, which is consistent with its standing as the telecommunication industry’s gold-standard standard for usage-based billing. I then use the hourly DPI data to calculate the proportion of traffic within each hour that is generated by different applications or protocols. These proportions multiplied by the hourly IPDR byte counts yield hourly byte counts for each application, which I use throughout my analysis.

DPI data records usage at the application level each hour. Given the vast number of applications, it is necessary to develop a taxonomy to group the applications by utility or function. I follow industry standards, and use a grouping very similar to that used by Sandvine in their regular reporting on aggregate Internet traffic. I present this classification in Table 1.1. Only applications in the top 175 by either total usage or frequency in the sample are assigned to a group with all remaining applications labeled Other. This simple grouping captures almost 99.9% of all traffic, with the other category including only 0.13% of traffic.

As a percentage of all subscriber usage, which is reported in Table 1.1, Video and Browsing are the two most popular online activities, together accounting for over 80% of traffic. While other activities such as Music and Gaming are commonplace, they comprise much smaller percentages of total traffic. In fact, outside of Video and Browsing no other group accounts for more than 10%.

1.3 Overview of Online Activity

In this section, I present an overview of subscriber usage at the monthly and hourly levels of observation. Since there are multiple months of observation for each subscriber in my sample, I take within-subscriber means across months to calculate average monthly and hourly usage. By doing so, I present a more representative picture of each subscriber’s behavior.
1.3.1 Monthly Descriptive Statistics of Online Activity

My sample contains monthly usage statistics on about 43,000 subscribers from a single market. 76% of these subscribers receive both broadband and pay TV service from the ISP with the remaining subscribers only receiving broadband service. Table 1.2 contains summary statistics on these two groups of subscribers. I refer to subscribers with pay TV and broadband service as being *bundled* and those with just broadband service as *unbundled*.

Unbundled subscribers pay about $10 more per month for broadband service and receive the same average (advertised) downstream speed, 50 Mbps, compared to bundled subscribers. Bundled subscribers typically receive a discount for each service as part of the bundle, which explains the difference I observe in prices.

Unbundled subscribers are also much heavier Internet users with average monthly usage that is 97% greater than bundled subscribers. I also observe monthly usage for unbundled subscribers to be more asymmetric, which is explained by half of their traffic being *Video*. Since bundled subscribers are able to watch video using the ISP’s pay TV service, it is unsurprising that the average proportion of traffic that is OTTV, across all bundled subscribers, is 36%. This is substantially lower than the 51% for bundled subscribers.\(^5\) It follows that from the ISP’s perspective delivering video to a consumer linearly is more cost effective due to multicast technology. This differs to OTTV where consumers are able to watch video in an on-demand fashion.

*Video* and *Browsing* clearly dominate a subscriber’s monthly usage, as expected from the aggregate group distribution presented in Table 1.1. From Table 1.3, I see average monthly *Video* usage is 55 GB and average *Browsing* usage is 27 GB. Even though *Video* and *Browsing* are the two most popular online activities, *Video* stands alone with twice as much monthly usage. *Streaming* usage is third, at just 6 GB per month.

\(^5\)Note that these statistics substantially understates the total proportion of OTTV overall, because the heaviest of users use much more video. Thus when averaging the proportions across subscribers, it understates their impact on the overall traffic composition.
Table 1.2: *Descriptive Statistics of Subscribers*

<table>
<thead>
<tr>
<th></th>
<th>Has Pay TV</th>
<th>No Pay TV</th>
<th>Everyone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mbps (Down)</td>
<td>48.5</td>
<td>50.1</td>
<td>48.9</td>
</tr>
<tr>
<td>Broadband Price ($)</td>
<td>52.49</td>
<td>63.14</td>
<td>55.05</td>
</tr>
<tr>
<td>Monthly Usage (GB)</td>
<td>81.14</td>
<td>159.22</td>
<td>99.52</td>
</tr>
<tr>
<td>Downstream %</td>
<td>92.9%</td>
<td>91.6%</td>
<td>91.9%</td>
</tr>
<tr>
<td>% of Usage Video</td>
<td>35.7</td>
<td>50.7</td>
<td>39.2</td>
</tr>
<tr>
<td>N</td>
<td>32,876</td>
<td>10,124</td>
<td>43,000</td>
</tr>
</tbody>
</table>

*Note:* All values reported are averages at the subscriber level of observation. Average megabits per second are based on what is advertised with the subscriber’s broadband service tier. % of Usage Video represents the average percentage of monthly traffic from the Video group described in Table 1.1. Pay TV is considered to be linear video service that is sold by the ISP.

I find that many online activities such as *Gaming, Music,* and *Sharing,* while popular in the media and culturally, do not generate much usage comparatively. In fact, I observe in Table 1.3 very little usage by subscribers below the median, which suggests these activities are popular among a subset of subscribers. For example, the 75th percentile *Gaming* subscriber uses about 0.82 GB per month, or 58 times less than the 99th percentile. As reference, the 75th percentile *Video* subscriber only uses 5.6 times less than the 99th percentile.

It is worth noting that the biggest difference between *Video* and *Browsing* usage appears to come from the heaviest *Video* users. In Table 1.3 I observe similar medians between the two groups and a greater 25th percentile *Browsing* user than for *Video.* The two distributions diverge quickly with the 75th percentile *Video* user consuming over twice as much as *Browsing.* These results suggest *Browsing* usage is flatter across all subscribers, *Video* usage is more disperse, and that these two activities are unique in that almost every subscriber participates in them, unlike the smaller groups discussed above.

These similarities and differences between *Video* and *Browsing* indicate subscriber usage by group varies proportionally depending on how heavy the subscriber’s usage is. That is, is a greater proportion of usage *Video* for heavier Internet subscribers? I normalize traffic
Table 1.3: Average Monthly Usage by Group

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>Median</th>
<th>25th Pctile</th>
<th>75th Pctile</th>
<th>90th Pctile</th>
<th>99th Pctile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administration</td>
<td>1.19</td>
<td>0.27</td>
<td>0.07</td>
<td>0.89</td>
<td>2.65</td>
<td>15.12</td>
</tr>
<tr>
<td>Backup</td>
<td>0.58</td>
<td>0.05</td>
<td>0.00</td>
<td>0.28</td>
<td>0.85</td>
<td>7.01</td>
</tr>
<tr>
<td>Browsing</td>
<td>26.57</td>
<td>14.95</td>
<td>6.35</td>
<td>30.90</td>
<td>57.88</td>
<td>196.74</td>
</tr>
<tr>
<td>CDN</td>
<td>2.93</td>
<td>0.63</td>
<td>0.10</td>
<td>2.41</td>
<td>6.83</td>
<td>35.46</td>
</tr>
<tr>
<td>Gaming</td>
<td>3.05</td>
<td>0.01</td>
<td>0.00</td>
<td>0.82</td>
<td>8.22</td>
<td>47.32</td>
</tr>
<tr>
<td>Music</td>
<td>3.38</td>
<td>1.37</td>
<td>0.17</td>
<td>4.16</td>
<td>8.60</td>
<td>26.42</td>
</tr>
<tr>
<td>Sharing</td>
<td>0.20</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td>0.00</td>
<td>0.20</td>
</tr>
<tr>
<td>Streaming</td>
<td>6.23</td>
<td>1.36</td>
<td>0.35</td>
<td>4.50</td>
<td>12.58</td>
<td>89.12</td>
</tr>
<tr>
<td>Tunnel</td>
<td>0.07</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.20</td>
</tr>
<tr>
<td>Video</td>
<td>55.20</td>
<td>18.29</td>
<td>1.77</td>
<td>72.97</td>
<td>159.56</td>
<td>405.75</td>
</tr>
<tr>
<td>Other</td>
<td>0.13</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.13</td>
<td>2.19</td>
</tr>
<tr>
<td>All</td>
<td>99.52</td>
<td>50.81</td>
<td>14.31</td>
<td>136.39</td>
<td>257.57</td>
<td>596.73</td>
</tr>
</tbody>
</table>

Note: This table reports monthly summary statistics in gigabytes (GB) by group for 43,000 subscribers between April 2015 and August 2015. All summary statistics reported are relative to each group and are not representative of the median subscriber, for example, for the entire sample. That is, there is no relationship across rows. The All row reports statistics on aggregate usage for a subscriber across all groups.

for each quantile of the total usage distribution in panel (a) of Figure 1.1 to answer this question. In the figure, the proportion of quantile usage that is Video, Browsing, Music & Streaming, and Other, which contains all other groups in Table 1.1, is presented. I find the proportion of usage that is Video increases and Browsing decreases among heavier Internet users. Moreover, the proportion of traffic spent on Music & Streaming, and Other remains constant around 20%.

Usage levels by quantile, which are presented in panel (b) of Figure 1.1, clearly shows how Video is driving the majority of heavy usage observed. In the past, much attention was paid to activities such as file sharing and gaming as culprits of driving the heavy usage of “super users”. While these activities may have been more notable historically, I see no evidence of these types of activities driving usage of the heaviest subscribers. For example, the 99th
(a) Proportions of Total Usage

Note: Panel (a) reports the proportion of total usage by quantile that is Video, Browsing, Music/Streaming, and Other, where other includes all groups that are not Video, Browsing, Music, or Streaming. Panel (b) reports average usage by quartile for the same four groups.

percentile subscriber from panel (b) watches around 380 GB of Video per month compared to just 20 GB for the median subscriber and 148 GB for the 90th percentile subscriber, respectively.

1.3.2 Hourly Descriptive Statistics of Online Activity

I plot average hourly usage in Figure 1.2 in aggregate and decomposed into Video, Browsing, Music & Streaming, and Other. The observed peak-trough pattern in panel (a) is unchanged from previous research: hourly usage is lowest during early morning hours and highest during the late evening. Average hourly usage peaks during the 10PM hour at just over 250 MB/hour with a nadir of 50 MB/hour during the 5AM hour. I also observe more downstream than upstream usage across the entire day, which is consistent with OTTV being the majority of traffic.
Figure 1.2: *Average Hourly Usage*

![Graph](image)

(a) Downstream/Upstream Usage  
(b) Decomposed Hourly Usage

**Note:** Panel (a) reports average downstream and upstream usage in megabytes (MB) by hour for a day. Panel (b) reports average usage decomposed into Video, Browsing, Music & Streaming, and Other, where other includes all groups that are not Video, Browsing, Music, or Streaming.

Most importantly, usage for each group in panel (b) of Figure 1.2 follows the same peak-trough pattern observed in the aggregate of panel (a). This suggests the costs of delivering different types of traffic are proportional to the type’s overall usage. That is, Video, in this case, is the most expensive online activity since it generates the largest proportion of usage in the evening. This result complements the findings of Malone et al. (2014) where they observe subscribers across the total usage distribution spending roughly the same proportion of usage during the evening.

When moving from daytime to evening hours, I observe in Table 1.4 hourly usage more than doubling for deciles 4 through 9. The tails of the distribution increase too, but to a lesser degree. Similar to the findings in Malone et al. (2014), the fact hourly usage increases in the evening for the entire distribution implies the cost of serving a subscriber is indeed proportional to her usage. Moreover, since operators invest in the network to handle evening
Table 1.4: *Average Hourly Total Usage During Off-Peak and Peak Hours*

<table>
<thead>
<tr>
<th>Decile</th>
<th>12:00AM−5:59PM</th>
<th>6:00PM−11:59PM</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.18</td>
<td>4.64</td>
<td>1.46</td>
</tr>
<tr>
<td>2</td>
<td>9.45</td>
<td>16.51</td>
<td>1.75</td>
</tr>
<tr>
<td>3</td>
<td>18.01</td>
<td>33.77</td>
<td>1.87</td>
</tr>
<tr>
<td>4</td>
<td>29.86</td>
<td>61.26</td>
<td>2.05</td>
</tr>
<tr>
<td>5</td>
<td>46.57</td>
<td>102.63</td>
<td>2.20</td>
</tr>
<tr>
<td>6</td>
<td>71.04</td>
<td>158.43</td>
<td>2.23</td>
</tr>
<tr>
<td>7</td>
<td>103.77</td>
<td>233.46</td>
<td>2.25</td>
</tr>
<tr>
<td>8</td>
<td>154.13</td>
<td>335.20</td>
<td>2.17</td>
</tr>
<tr>
<td>9</td>
<td>235.29</td>
<td>483.81</td>
<td>2.06</td>
</tr>
<tr>
<td>10</td>
<td>482.61</td>
<td>871.33</td>
<td>1.81</td>
</tr>
</tbody>
</table>

*Note:* This table reports average hourly usage by peak and off-peak hours in megabytes (MB). These averages include traffic from all groups. The *Ratio* column reports the ratio of peak:off-peak usage.

Traffic, these results also suggest any technologies that enable the delivery of content to daytime hours would help improve efficiency by freeing bandwidth in the evening.

Hourly *Video* usage changes the most between daytime and evening hours. In Table 1.5, I observe hourly *Video* usage more than doubling from 12:00AM−5:59PM to 6:00PM−11:59PM, except for the 1<sup>st</sup> and 10<sup>th</sup> deciles of users. For example, average hourly *Video* usage is 2.8 times larger in the evening for the 5<sup>th</sup> decile of users. While usage for *Browsing, Music & Streaming*, and *Other* all increase in the evening, none do so at the rate of *Video*.

Increased usage in the evening is partly explained by more members of a family being home after work and/or school. Obviously, if more people are using an Internet connection, usage will increase. However, *Video* being the most pronounced may be suggestive of other market changes. First, greater *Video* usage is indicative of consumers substituting for pay TV with OTTV. Second, it signals *Video* becoming a more engrained aspect of daily life, where downtime is now filled with YouTube or Netflix video.
Table 1.5: Average Hourly Group Usage During Off-Peak and Peak Hours

<table>
<thead>
<tr>
<th>Decile</th>
<th>12:00AM–5:59PM</th>
<th></th>
<th></th>
<th>6:00PM–11:59PM</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Video</td>
<td>Browsing</td>
<td>Music &amp; Stream.</td>
<td>Other</td>
<td>Video</td>
<td>Browsing</td>
</tr>
<tr>
<td>1</td>
<td>0.37</td>
<td>2.22</td>
<td>0.29</td>
<td>0.31</td>
<td>0.67</td>
<td>3.05</td>
</tr>
<tr>
<td>2</td>
<td>1.55</td>
<td>6.06</td>
<td>1.11</td>
<td>0.73</td>
<td>3.85</td>
<td>9.60</td>
</tr>
<tr>
<td>3</td>
<td>3.77</td>
<td>10.39</td>
<td>2.42</td>
<td>1.44</td>
<td>9.52</td>
<td>17.81</td>
</tr>
<tr>
<td>4</td>
<td>8.05</td>
<td>15.34</td>
<td>3.93</td>
<td>2.53</td>
<td>22.41</td>
<td>26.83</td>
</tr>
<tr>
<td>5</td>
<td>16.80</td>
<td>19.84</td>
<td>5.78</td>
<td>4.16</td>
<td>46.92</td>
<td>36.75</td>
</tr>
<tr>
<td>6</td>
<td>31.01</td>
<td>25.27</td>
<td>8.49</td>
<td>6.28</td>
<td>82.67</td>
<td>47.23</td>
</tr>
<tr>
<td>7</td>
<td>51.26</td>
<td>31.79</td>
<td>11.55</td>
<td>9.18</td>
<td>131.85</td>
<td>60.92</td>
</tr>
<tr>
<td>8</td>
<td>84.89</td>
<td>41.21</td>
<td>15.01</td>
<td>13.01</td>
<td>202.15</td>
<td>78.14</td>
</tr>
<tr>
<td>9</td>
<td>137.12</td>
<td>56.99</td>
<td>21.53</td>
<td>19.66</td>
<td>301.90</td>
<td>102.93</td>
</tr>
<tr>
<td>10</td>
<td>283.34</td>
<td>115.01</td>
<td>44.40</td>
<td>39.86</td>
<td>540.62</td>
<td>184.46</td>
</tr>
</tbody>
</table>

Note: This table reports average hourly usage in megabytes (MB) during Peak (6:00PM–11:59PM) and Off-Peak (12:00AM–5:59PM) hours by decile. The deciles reported are based on total monthly usage across all groups. Other includes all groups that are not Video, Browsing, Music, or Streaming.

Proportionally, subscribers tend to trade-off Browsing for Video, with the largest changes coming from the middle of the distribution. For example, in Table 1.6, I observe Video accounts for a larger proportion of 6:00PM–11:59PM usage than during the rest of the day for all deciles. Moreover, Browsing usage goes down for each decile and the proportion of Music & Streaming and Other traffic remains constant. That is, the only notable changes I see are between Video and Browsing. For the 4th and 5th deciles, this is as much as an additional 10% of usage going to Video, whereas it is only 3% for the 10th decile. However, percentage changes in usage may be somewhat misleading in this case given the differences in usage levels observed in Table 1.5.

1.3.3 Persistence of Group Usage

I calculate the persistence of Video, Browsing, Music & Streaming, and Other usage by calculating monthly deciles for each group and totaling the month-to-month transitions be-
### Table 1.6: Normalized Usage for Peak and Off-Peak Hours

<table>
<thead>
<tr>
<th>Decile</th>
<th>12:00AM–5:59PM</th>
<th>6:00PM–11:59PM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Video</td>
<td>Browsing</td>
</tr>
<tr>
<td>1</td>
<td>0.12</td>
<td>0.70</td>
</tr>
<tr>
<td>2</td>
<td>0.16</td>
<td>0.64</td>
</tr>
<tr>
<td>3</td>
<td>0.21</td>
<td>0.58</td>
</tr>
<tr>
<td>4</td>
<td>0.27</td>
<td>0.51</td>
</tr>
<tr>
<td>5</td>
<td>0.36</td>
<td>0.43</td>
</tr>
<tr>
<td>6</td>
<td>0.44</td>
<td>0.36</td>
</tr>
<tr>
<td>7</td>
<td>0.49</td>
<td>0.31</td>
</tr>
<tr>
<td>8</td>
<td>0.55</td>
<td>0.27</td>
</tr>
<tr>
<td>9</td>
<td>0.58</td>
<td>0.24</td>
</tr>
<tr>
<td>10</td>
<td>0.59</td>
<td>0.24</td>
</tr>
</tbody>
</table>

*Note:* This table reports the proportion of all traffic during Peak (6:00PM–11:59PM) and Off-Peak (12:00AM–5:59PM) usage by group and decile. The deciles reported are based on total monthly usage across all groups. For both Peak and Off-Peak hours, the proportions sum to 1, respectively. Other includes all groups that are not Video, Browsing, Music, or Streaming.

tween these deciles. These transition matrices are presented as proportions graphically in Figure 1.3.

Calculating the persistence of usage in this manner is similar to the approach taken in Malone et al. (2014); however, the authors have no details on traffic composition. They find monthly usage is persistent with over 90% of month-to-month transitions occurring within ±2 deciles and the tails of the usage distribution the most persistent.

The saddle shape of the surfaces depicted in Figure 1.3 show subscriber usage is persistent within groups and that the tails of each group distribution are the most persistent. For example, 70% of subscribers in the 10th decile of Video one month are in the 10th decile of Video the following month. These findings are consistent across all four groups presented in Figure 1.3 suggesting subscribers tend to enjoy the same activities month-to-month.

---

<sup>6</sup>Calculating these transitions require a subscriber to be present at least two months in the sample.
In Sections 1.3.1 and 1.3.2, the larger magnitude of Video usage compared to other types of online activity is clear. From Figure 1.3, I observe strong persistence in the 10th decile of Video subscribers. Together, these results indicate the heaviest Video users consistently generate a disproportionate amount of the costs on the network. If no persistence was observed in Figure 1.3 then the high costs of Video usage would be shared more evenly across all subscribers.

More generally, the persistence in Figure 1.3 means subscribers tend to follow certain subscriber types since they do the same things online in similar volume. That is, if a subscriber tends to be a heavy Browsing user one month, she is likely to be a heavy Browsing user the next month. The same holds true for all of the groups listed in Table 1.1. This is consistent with the persistent type assumption made in Nevo et al. (2016).

1.4 Cord Cutting and Pay TV Substitution

In Section 1.3, the important role Video usage plays in driving heavy Internet usage is clear, accounting for 55% of all traffic. As such, any market changes that leads to future increases in Video usage are relevant. Cord cutting is one such phenomenon that would lead to significant increase in Video usage in the future. A subscriber that cuts the cord cancels all pay TV service through the operator but remains a broadband customer. It follows these subscribers must turn to other alternatives for video services with an increasingly compelling option now being OTTV.\footnote{Cord cutters could also turn to satellite service as a pay TV, mobile, or over-the-air (OTA) alternatives.}

OTTV and pay TV were once thought to be complementary, but recent trends in the video market suggest otherwise. For instance, the number of North American households are increasing, but pay TV penetration has failed to grow, which suggest more households are substituting OTTV for pay TV. Many OTTV options such as Netflix, Hulu, and Amazon
Table 1.7: Percentage of Cord Cutters by Decile

<table>
<thead>
<tr>
<th>Decile</th>
<th>Pct of Cord Cutters</th>
<th>Pct of Cord Cutters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total Usage</td>
<td>Video Usage</td>
</tr>
<tr>
<td></td>
<td>Ranking</td>
<td>Ranking</td>
</tr>
<tr>
<td>1</td>
<td>2.5</td>
<td>3.2</td>
</tr>
<tr>
<td>2</td>
<td>5.3</td>
<td>4.6</td>
</tr>
<tr>
<td>3</td>
<td>6.0</td>
<td>6.3</td>
</tr>
<tr>
<td>4</td>
<td>7.4</td>
<td>3.9</td>
</tr>
<tr>
<td>5</td>
<td>9.2</td>
<td>8.5</td>
</tr>
<tr>
<td>6</td>
<td>9.5</td>
<td>12.0</td>
</tr>
<tr>
<td>7</td>
<td>13.7</td>
<td>11.6</td>
</tr>
<tr>
<td>8</td>
<td>11.6</td>
<td>13.0</td>
</tr>
<tr>
<td>9</td>
<td>14.1</td>
<td>16.5</td>
</tr>
<tr>
<td>10</td>
<td>20.8</td>
<td>20.4</td>
</tr>
</tbody>
</table>

Note: This table reports the percentage of the 283 “cord cutters” by deciles of all usage and all Video usage. Each column of this table sums to 100.

offer on-demand libraries of content, limited ads, and a portability at lower price points that are attractive to certain segments of the population. If OTTV and pay TV are, in fact, substitutes, as more subscribers cord cut, the volume of Video traffic on the network will increase.

Out of the 43,000 subscribers in my sample, I observe 283 subscribers who cut the cord. The vast majority of these cord cutters, as reported in Table 1.7, are heavy Internet users with 47% coming from the top 30% of Internet users and 50% coming from the top 30% of Video users. Given the strong correlation between heavy Internet and Video usage, these results are unsurprising. A heavy Video user will be most familiar with OTTV options making a transition away from pay TV reasonable.

I estimate the effect of cord cutting on subscriber behavior by calculating the difference in average daily usage before and after the drop of pay TV service.\textsuperscript{8} In Table 1.8, daily

\textsuperscript{8}These calculations use all data before and after the date I observe a subscriber cut the cord. I repeated the analysis with varying windows on either side and the results do not change.
Table 1.8: Unconditional Changes in Average Daily Usage by Group

<table>
<thead>
<tr>
<th>Group</th>
<th>Before</th>
<th>After</th>
<th>Diff</th>
<th>% Change</th>
<th>% of Total Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Browsing</td>
<td>1.08</td>
<td>1.30</td>
<td>0.22</td>
<td>20.14</td>
<td>16.79</td>
</tr>
<tr>
<td>Video</td>
<td>2.76</td>
<td>3.68</td>
<td>0.92</td>
<td>33.38</td>
<td>70.23</td>
</tr>
<tr>
<td>Music &amp; Streaming</td>
<td>0.44</td>
<td>0.48</td>
<td>0.04</td>
<td>8.17</td>
<td>3.05</td>
</tr>
<tr>
<td>Other</td>
<td>0.32</td>
<td>0.46</td>
<td>0.13</td>
<td>40.57</td>
<td>9.92</td>
</tr>
<tr>
<td>All Groups</td>
<td>4.61</td>
<td>5.92</td>
<td>1.31</td>
<td>28.35</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Note: This table reports changes in daily usage for cord cutters in my sample. All averages are reported in gigabytes (GB) and are unconditional. That is, all 283 cord cutters are included. Before is average usage before dropping pay TV, After is average usage after drop pay TV, % Change is the % change in row usage after dropping pay TV, and % of Total Change reports the percentage of change in all usage (1.31 GB) the row accounts for. The % of Total Change sums to 100%.

Video usage increases by 33%, which aside from Other, is the largest percentage increase. The additional 0.92 GB of Video per day is also the largest difference in levels, too. Using a rough conversion of 1.5 GB per hour of HD streaming this increase in Video is equal to an additional 36 minutes of OTTV video per day.

I also observe economically significant increases in Browsing usage, but to a lesser degree than with Video. Music & Streaming show little change, and the magnitude of Other is small. There is little difference between the conditional and unconditional changes in daily usage since essentially all subscribers have positive usage for these four groups.

The % of Total Change column of Table 1.8 takes the difference in daily usage for each row and reports its part of the 1.31 GB increase in daily usage. I find Video accounts for 70% of the increase in daily usage with Browsing the next largest contributor at 17%. Music & Streaming and Other both increase, but the relative magnitudes when compared to Video and Browsing are small, so the overall contribution pales in comparison.

---

9This assumes an average stream rate of 3.33 Mbps. A lower bit rate assumption would imply greater usage.
While the large increase in Video usage observed in Table 1.8 suggests subscribers are substituting for pay TV with OTTV, a clearer answer can be made if I look for large changes in daily usage by video type. In Table 1.9 I break out Video usage into many popular OTTV options such as Netflix, Hulu, YouTube, and Amazon. Just as in Table 1.8, I report unconditional usage.

The video types with the largest increases in daily usage tend to be those with libraries and features more comparable to pay TV: Netflix, Hulu, and Sling TV, for example. However, YouTube and Flash video see hardly any change in daily usage. YouTube and Flash videos tend to be user generated and of a different type than is offered on Netflix, for example. These results strongly suggest the increase in Video usage observed in Table 1.8 is from cord cutters substituting for pay TV with OTTV.

When accounting for the 0.92 GB/day increase in Video usage, Netflix is by far the largest factor explaining 75% of this increase. Sling TV and Hulu are second and third, respectively. As discussed above, these three options tend to be the most comparable to a traditional pay TV offering and are likely to be bundled together to substitute for pay TV.

Table 1.10 reports the same information as in Table 1.9 but conditions on positive usage for each video type. That is, only cord cutters with positive Netflix usage in my sample are included in the Netflix usage statistics. Since YouTube is used by almost everyone, I see little change. However, for OTTV services such as Sling TV, Hulu, and HBO GO, I observe large shifts in daily usage. I present these conditional averages to emphasize that at an aggregate level the impact of different OTTV services may be underestimated. Netflix usage is still the dominant Video option. If I use the same GB-to-HD video conversion from above, cord cutters who use Netflix watch about 2.25 hours of HD video a day just from Netflix, ignoring any other OTTV services they may subscribe to.
Table 1.9: Unconditional Changes in Average Daily Usage by Video Type

<table>
<thead>
<tr>
<th>Type</th>
<th>Before</th>
<th>After</th>
<th>Diff</th>
<th>% Change</th>
<th>% of Total Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon</td>
<td>0.05</td>
<td>0.03</td>
<td>-0.02</td>
<td>-48.07</td>
<td>-2.17</td>
</tr>
<tr>
<td>Flash</td>
<td>0.02</td>
<td>0.03</td>
<td>0.00</td>
<td>12.83</td>
<td>0.00</td>
</tr>
<tr>
<td>HBO Go</td>
<td>0.03</td>
<td>0.05</td>
<td>0.02</td>
<td>47.07</td>
<td>2.17</td>
</tr>
<tr>
<td>Hulu</td>
<td>0.09</td>
<td>0.15</td>
<td>0.06</td>
<td>72.44</td>
<td>6.52</td>
</tr>
<tr>
<td>Netflix</td>
<td>1.91</td>
<td>2.60</td>
<td>0.69</td>
<td>36.30</td>
<td>75.00</td>
</tr>
<tr>
<td>Sling TV</td>
<td>0.01</td>
<td>0.10</td>
<td>0.09</td>
<td>1044.98</td>
<td>9.78</td>
</tr>
<tr>
<td>YouTube</td>
<td>0.55</td>
<td>0.63</td>
<td>0.07</td>
<td>13.34</td>
<td>7.61</td>
</tr>
<tr>
<td>Other</td>
<td>0.09</td>
<td>0.10</td>
<td>0.00</td>
<td>4.32</td>
<td>0.00</td>
</tr>
<tr>
<td>All Video</td>
<td>2.76</td>
<td>3.68</td>
<td>0.92</td>
<td>33.38</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Note: This table reports changes in daily usage for cord cutters in my sample. All averages are reported in gigabytes (GB) and are unconditional. That is, all 283 cord cutters are included. Before is average usage before dropping pay TV, After is average usage after drop pay TV, % Change is the % change in row usage after dropping pay TV, and % of Total Change reports the percentage of change in all usage (1.31 GB) the row accounts for. The % of Total Change sums to 100%.

1.5 Conclusion

I analyze high-frequency disaggregated data on the volume and composition of households’ Internet usage to provide insights into a number of ongoing public policy debates. I find that nearly two thirds of peak Internet usage is OTTV, and that nearly every household engages OTTV to some degree. Dramatic differences in the level of peak usage across subscribers, the determinant of network costs imposed on the ISP by a subscriber, are largely due to differences in the level of OTTV usage. I also find that temporal usage patterns across the day are nearly identical regardless of the total level of usage by a subscriber. The level and composition of household usage for different applications, and overall, are also highly persistent and predictable at both the daily and monthly level.

Perhaps my most interesting results are the dramatic difference between the usage behavior of subscribers with and without traditional TV service. Broadband subscribers that
Table 1.10: Conditional Changes in Average Daily Usage by Video Type

<table>
<thead>
<tr>
<th>Type</th>
<th>Before</th>
<th>After</th>
<th>Diff</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon</td>
<td>0.29</td>
<td>0.17</td>
<td>-0.12</td>
<td>-42.30</td>
</tr>
<tr>
<td>Flash</td>
<td>0.03</td>
<td>0.04</td>
<td>0.00</td>
<td>11.77</td>
</tr>
<tr>
<td>HBO Go</td>
<td>0.21</td>
<td>0.29</td>
<td>0.08</td>
<td>35.30</td>
</tr>
<tr>
<td>Hulu</td>
<td>0.31</td>
<td>0.43</td>
<td>0.11</td>
<td>36.22</td>
</tr>
<tr>
<td>Netflix</td>
<td>2.66</td>
<td>3.39</td>
<td>0.73</td>
<td>27.51</td>
</tr>
<tr>
<td>Sling TV</td>
<td>0.42</td>
<td>0.87</td>
<td>0.45</td>
<td>108.18</td>
</tr>
<tr>
<td>YouTube</td>
<td>0.56</td>
<td>0.63</td>
<td>0.08</td>
<td>13.74</td>
</tr>
<tr>
<td>Other</td>
<td>0.10</td>
<td>0.10</td>
<td>0.00</td>
<td>2.72</td>
</tr>
</tbody>
</table>

Note: This table reports changes in daily usage for cord cutters in my sample. All averages are reported in gigabytes (GB) and are conditional. That is, only subscribers with positive usage for a video type are included in these calculations. Before is average usage before dropping pay TV, After is average usage after drop pay TV, and % Change is the % change in row usage after dropping pay TV.

...
efficient pricing plans that introduce incentives to shift usage to off-peak hours rather than simple three-part tariffs, I may also observe changes from OTTV providers that will yield substantial welfare improvements for all parties. Features like Amazon’s ASAP can easily be easily adapted to utilize off-peak hours to locally cache content that the subscriber is likely to watch, and Netflix’s release of information that suggests a strong persistence in subscribers’ viewing habits, suggest that future usage can be almost perfectly predicted. Together, the technological advances by OTTV providers and the predictability of viewing habits suggest that very mild forms of peak-use pricing could be quite effective at lowering the network costs of ISPs while simultaneously improving the performance of many OTTV services.

My results also suggest that there is likely to be an accelerated transition away from traditional bundling practices towards “skinny bundles”. In my data, I observe that Sling TV, which is essentially a small bundle of channels including ESPN, is one of the most widely adopted OTTV services after a broadband subscriber drops traditional pay TV. Since the ISP loses a pay TV customer and experiences a large increase in network costs after a subscriber cuts the cord, I expect that skinny bundles may soon be offered to actual or potential cord cutters at a discounted rate in the very near future.

Finally, even though my results have potentially less direct implications for the net-neutrality debate and ongoing merger reviews in the telecommunication industry, I am confident they will serve as important inputs into future theoretical work on these topics. I believe my descriptive analysis is also an important first step in motivating further empirical research into various aspects of demand for residential broadband that I highlighted here. I look forward to progress in both areas as the telecommunication industry continues to undergo rapid changes.
Figure 1.3: Transition Matrices of Group Usage

(a) Video

(b) Browsing

(c) Music and Streaming

(d) Other

Note: These figures represent monthly transitions between deciles by group. These transition matrices are generated by calculating each subscriber’s monthly decile of usage and then tracking what decile they are in the next month. For example, from panel (a), about 70% of subscribers in the 10th decile of Video are in the 10th decile of Video the next month.
Chapter 2

Do Three-Part Tariffs Improve Efficiency in Residential Broadband Networks?

2.1 Introduction

The Internet is now an integral part of modern life. Recent research estimates the average American spends over 3 hours online daily.\(^1\) Moreover, with the proliferation of faster access speeds and bandwidth intensive applications like online video, which now comprise over 60% of peak usage, subscribers use more data than ever before.\(^2\) This growth in subscriber demand puts pressure on Internet Service Providers (ISPs) to manage demands placed on their networks.

In response, a sizable number of US providers now sell service via usage-based pricing plans. The most typical “three-part tariff” plan specifies an access fee, usage allowance and

an overage price. Subscribers who use less data than the allowance pay just the access fee for service, while subscribers who use more pay the overage price for each additional GB used.\footnote{Some plans bill for overage more crudely. For example, Comcast’s XFINITY plan bills $10 for additional 50 GB blocks of data, and also gives subscribers three warnings before imposing the fees.}

ISPs typically argue that such plans lower overall and peak usage, helping to reduce network costs, and reduce the level of cross-subsidization between light and heavy users, more closely linking subscriber costs to usage. Government agencies such as the US Federal Communications Commission (FCC) are closely watching to see how usage-based pricing affects costs and efficiency (OIAC (2013)). Yet, there is little empirical evidence to date.

In this paper, I analyze subscriber-specific usage data from a North American ISP during May 2011–May 2013. Importantly, this provider sells service via a menu of three-part tariff plans but also sells service via unlimited plans to a group of grandfathered customers. In comparing usage behavior across groups of subscribers, I identify key effects of three-part tariff pricing. I also analyze how subscriber usage changes across the day and across the month for a single billing cycle.

First, I analyze monthly data. Despite enjoying faster connection speeds, the top ten (one) percent of users on unlimited plans use about 56% (74%) more data in May 2013 than the top ten (one) percent of users facing three-part tariffs. In contrast, the median user who faces a three-part tariff uses just 10% less data than the median user on an unlimited plan. Other than the top few deciles of users, distributions of usage are quite similar across subscriber groups. Thus, allowances and overage prices are mostly effective at reining in heavy users. Since such users account for a disproportionate share of total usage, three-part tariffs save significant network costs for the ISP.

By linking payments to usage, three-part tariffs also reduce the amount of cross-subsidization between the low- and high-volume users. Among subscribers on unlimited plans, the aver-
age subscriber in the top ten percent of users consumes 21.5 times more data per month but pays just 1.1 times more than the average subscriber in the bottom half of the usage distribution. Among subscribers facing three-part tariffs, however, the average subscriber in the top ten percent of users consumes just 15 times more data per month than the average bottom-half-usage subscriber, but pays about 2.3 times more for service.

Three-part tariffs do not strongly affect how much subscribers favor downloading (e.g., streaming Netflix) versus uploading (e.g., moving a file from local storage to Dropbox). Subscribers facing three-part tariffs download an average of 90.3% of data per month. While this is statistically smaller than the 90.9% fraction for subscribers to unlimited plans, I do not view a 0.6 percentage point difference as economically significant.

Second, I more closely examine data aggregated up to daily peak hours and daily off-peak hours for a one-month billing cycle during May and June of 2012. I find that subscribers facing three-part tariffs change their usage throughout the month in ways consistent with forward-looking behavior. Subscribers on pace to exceed their allowance cut back on usage in statistically and economically significant ways. For example, subscribers with cumulative usage at between 80% and 100% of the allowance face high implicit overage prices and reduce usage by 28% over the last six days of the billing cycle. In contrast, those who have consumed under 40% of the usage allowance face low implicit overage prices and increase usage by 13% over the last six days of the month.

Most importantly, the percentage effects described above—indeed, all effects of cumulative usage and time of month on current usage—are *statistically the same during peak and off-peak hours*. Intuitively, the way that subscribers cut back does not depend on time of day. Hence, while three-part tariffs do curtail overall usage by heavy users, these users lower their peak usage by the same proportion as they lower their off-peak usage.

Because total off-peak usage is (by definition) below network capacity constraints, any reduction in overall usage that occurs during off-peak hours has no effect on ISP welfare.
but lowers welfare for both subscribers and content providers (e.g., Netflix), who lose value-creating transactions. This suggests service plans that differentially price peak and off-peak usage may improve upon the three-part tariff schedule’s ability to enhance consumer welfare and increase ISP profitability. Such plans would also provide incentives for content providers to make content more portable across the day. For example, since Netflix offers no “live” content, much of this traffic (currently over 30% of peak traffic) could be downloaded during off-peak hours.

I also show that usage is highly persistent, in the sense that a subscriber’s usage in one month almost perfectly predicts usage in the next month. In the monthly data, more than 90 percent of subscribers use an amount that places them within two deciles of their position in the prior month’s usage distribution. In the intra-month data, over half of the variation in usage is explained by individual-specific fixed effects. Hence, heavy users in one period tend to be heavy users in the next period. This is important, because it makes it relatively easy for ISPs to target such users with three-part tariffs while limiting the impact to a small number of subscribers.

My study contributes to the policy debate surrounding appropriate network management. Due to uncertainty about how government agencies such as the FCC will enforce net-neutrality, ISPs have been reluctant to negotiate contracts with content providers to pay for the costs of delivering their traffic to end users. However, usage-based pricing of subscriber service plans appears to be (for the moment) a safe harbor from such scrutiny. For example, the 2010 FCC Open Internet Report and Order states:

“... prohibiting tiered or usage-based pricing and requiring all subscribers to pay the same amount for broadband service, regardless of the performance or usage of the service, would force lighter end users of the network to subsidize heavier end users. It would also foreclose practices that may appropriately align incentives to encourage efficient use of networks. The framework we adopt today does not
prevent broadband providers from asking subscribers who use the network less to pay less, and subscribers who use the network more to pay more.” (FCC (2010), Paragraph 72)

While three-part tariff pricing is common practice by U.S. cellular providers and foreign residential broadband companies, it and other forms of usage-based pricing are controversial in the United States. Numerous consumer groups argue usage-based pricing is unfair and an unnecessary use of market power,⁴ and Senator Ron Wyden has proposed legislation restricting it.⁵ To inform this debate, the Federal Communications Commission tasked its Open Internet Advisory Committee to study the economics of usage-based pricing. The committee’s recent report (OIAC (2013)) includes information on the proliferation of usage-based pricing and sets up a useful framework for thinking about the efficiency of usage-based pricing and its potential effects on access. However, the report has limited access to usage data and leaves most questions unanswered.

Beyond policy insights, my analysis contributes to multiple literatures. The first studies access to and demand for broadband usage. A number of theoretical studies consider usage-based pricing (Mackie-Mason and Varian (1995); Bauer and Wildman (2012); Odlyzko et al. (2012)), primarily to focus on its welfare implications. On the empirical side, Hitt and Tambe (2007) show that broadband adoption significantly increases usage, while several papers estimate the economic value of broadband internet (Dutz et al. (2009); Rosston et al. (2010); Greenstein and McDevitt (2011)). A small number of papers use highly-detailed and disaggregated usage data. Edell and Varaiya (2002) and Varian (2002) use data from the INDEX experiments, where subjects were offered different prices for different (and very slow, by today’s standards) broadband speeds, to estimate demand for broadband and opportunity

---

⁴See, e.g., StopTheCap.com.

A separate literature examines whether consumers are forward-looking and able to make rational decisions when faced with complicated pricing schedules. These studies span a wide range of topics from telecommunications (Miravete (2003); Grubb and Osborne (2012)), to stockpiling behavior in response to sales (Hendel and Nevo 2006), to optimally using health insurance (Aron-Dine et al. (2012); Handel (2013)).

The remainder of the paper is as follows. Section 2.2 describes the data in detail and presents statistics on plan features. Section 2.3 then provides descriptive statistics on subscriber usage at the monthly and hourly levels. I test for how subscribers respond to changes in implicit prices of usage, and for how such responses depend on peak or off-peak hours, in Section 2.4. Section 2.5 concludes.

2.2 Data

The data used in this paper are from a North American ISP and contain information on subscriber usage from May 2011 to May 2013. All usage data originate from fifteen-minute level Internet Protocol Detail Records (IPDR), a trusted data source used for subscriber billing. I analyze data at a fifteen-minute frequency for May 12, 2012 to June 30, 2012. Outside of this window, I have access to monthly aggregates of a subscriber’s downstream and

---

6Due to data corruption issues, there are no data for March and April 2012.
upstream usage. I focus on this ISP’s largest markets to ensure the sample is representative of other North American communities.

IPDR records are collected by a Cable Modem Termination System (CMTS), which converts Internet traffic to a coaxial signal for home delivery. Figure 2.2 illustrates where a CMTS device is located on the network. The CMTS is marked by an oval.

My data also include variables on subscribers and plan choices. Each subscriber has a cable modem with a unique Media Access Control (MAC) address. Monthly plan-specific variables include a subscriber’s base price, usage allowance, and overage fees. I merge plan-specific variables with the IDPR data using the MAC address.

This ISP notably serves subscribers on both three-part tariff and unlimited usage plans. These unlimited plans are legacy options; new subscribers are offered just three-part tariff plans. I refer to subscribers remaining on unlimited plans as being grandfathered.

Table 2.2 presents statistics on plan details. From May 2011 to May 2013, the download speeds on three-part tariff plans almost double to 15.1 Mb/s, whereas speeds for unlimited
Table 2.1: *Summary Statistics of Plan Details*

<table>
<thead>
<tr>
<th></th>
<th>May 2011</th>
<th>May 2012</th>
<th>May 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Three-Part Tariff</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed Access Fee ($)</td>
<td>61.5</td>
<td>73.8</td>
<td>77.9</td>
</tr>
<tr>
<td>Per GB Overage Fee ($)</td>
<td>3.7</td>
<td>3.1</td>
<td>3.0</td>
</tr>
<tr>
<td>Down Speed (Mb/s)</td>
<td>8.9</td>
<td>13.7</td>
<td>15.1</td>
</tr>
<tr>
<td>Allowance Size (GB)</td>
<td>44.1</td>
<td>91.6</td>
<td>104.3</td>
</tr>
<tr>
<td>Over Allowance (%)</td>
<td>20.5</td>
<td>8.3</td>
<td>8.2</td>
</tr>
<tr>
<td>Allowance Used (%)</td>
<td>85.1</td>
<td>42.3</td>
<td>44.6</td>
</tr>
<tr>
<td>On Dominated Plan (%)</td>
<td>26.4</td>
<td>4.6</td>
<td>4.4</td>
</tr>
<tr>
<td>Subscribers</td>
<td>48,894</td>
<td>59,550</td>
<td>69,600</td>
</tr>
<tr>
<td><strong>Unlimited</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed Access Fee ($)</td>
<td>44.9</td>
<td>44.7</td>
<td>44.4</td>
</tr>
<tr>
<td>Per GB Overage Fee ($)</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Down Speed (Mb/s)</td>
<td>6.5</td>
<td>6.5</td>
<td>6.4</td>
</tr>
<tr>
<td>Allowance Size (GB)</td>
<td>∞</td>
<td>∞</td>
<td>∞</td>
</tr>
<tr>
<td>On Dominated Plan (%)</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Subscribers</td>
<td>28,075</td>
<td>17,426</td>
<td>11,761</td>
</tr>
</tbody>
</table>

**Note:** These statistics reflect plan characteristics and usage by subscribers to a single ISP during May 2011 - May 2013. Across plans, download speed is non-decreasing in the access fee and the overage price is non-increasing in the access fee. Usage is based upon Internet Protocol Detail Record (IPDR) data, captured in 15-minute intervals and aggregated to the monthly level. I say a plan is dominated if the subscriber could have chosen another plan and (holding usage constant) paid less and enjoyed advertised download speed no lower. Except for the count of subscribers, all reported values are averages at the subscriber level.

During the first year, the average usage allowance doubles in size to 91.6 GB/month. As a result, the percentage of people exceeding the usage allowance falls by 12% to 8.3%. Increasing the usage allowances also improves matching between subscribers and plans. Let a plan be “dominated” if there is another plan such that the subscriber would have paid less,
given their usage, and the advertised download speed is just as great. In May 2011, 26.4% of three-part tariff subscribers are on a dominated plan, but this falls to just 4.6% in May 2012, a drop of about 80%.

These statistics reveal that the average three-part tariff subscriber rarely exceeds the usage allowance. While this is partly due to higher usage allowances, subscribers are also better informed. This ISP now offers a notification system that alerts subscribers once common usage thresholds have been surpassed.\(^7\) For the entire panel, around 15% of subscriber-month observations exceed the usage allowance which translates to around 3.5 occurrences per subscriber over two years.

Subscribers who exceed the usage allowance nearly always use a consistently high amount over the entirety of the month. That is, overage charges rarely result from a small number of high-usage days. For subscribers who exceed the allowance in May 2012, only 0.3% of observations are for days where the subscriber uses more than 50% of the usage allowance in one day. This percentage increases to just 11.5% when 10% of the usage allowance is used as the cutoff. These results are further proof that subscribers are knowledgeable of the implications on overages by various Internet activities.

It is clear from Table 2.2 that there is significant movement between plans during 2011-13. Much of this movement is to higher-speed, higher-allowance plans. Figure 2.2 highlights how this reflects forward-looking behavior by subscribers. The Figure shows median growth rates and the percentage of subscribers who upgrade to plans with larger usage allowances, conditional on the portion of the usage allowance the subscriber consumed in May 2011 or 2012. Subscribers near the usage allowance experience the slowest growth and also show the highest rates of plan switching in the following year, which suggests subscribers are forward-looking and can effectively manage usage. Both ways in which these subscribers respond to

\(^7\)This is now mandated for the US cellular industry.
Figure 2.2: Yearly Growth and Plan Upgrade Incidence

Note: These statistics reflect usage and plan choices by subscribers to a single ISP during May 2011 - May 2013. Usage is based upon IPDR data, captured in 15-minute intervals and aggregated to the monthly level.

the possibility of overage fees are beneficial to the ISP. Slower growth rates save costs, and plan switching generates more revenue.

2.3 Monthly Usage

Differences among subscribers at the top of the usage distribution account for nearly all of the differences in average usage in May 2013. The cumulative distribution functions shown in Figure 2.3 are nearly identical at the median but diverge sharply for the top few deciles of usage. Table 2.3 shows that this divergence is statistically significant. A Kolmogorov-Smirnov (K-S) test, where the null hypothesis is that the distributions are equal, returns a highly significant test statistic of 0.074.8 By reining in these extreme subscribers, three-part tariffs also reduce the level of inequality. The Gini coefficient for usage among three-part tariff

---

8This statistic is well above the 1% critical value for this test, 0.006 (Massey (1951)). Results are qualitatively identical for May 2011 and May 2012.
Figure 2.3: Usage Distributions

Note: These cumulative distribution functions reflect usage by subscribers to a single ISP during May 2013. Usage is based upon IPDR data, captured in 15-minute intervals and aggregated to the monthly level.

Subscribers, 0.548, is nearly 9% lower than the Gini coefficient for usage among unlimited subscribers (0.599).\(^9\)

Despite these differences in the usage distributions, the composition of monthly usage is very similar among subscribers. Although a two-sample \(t\)-test rejects the null hypothesis that the downstream percentage of traffic is the same for three-part tariff and unlimited subscribers, the difference is not economically significant. For example, in May 2013, the unlimited subscribers’ average download percentage (90.9%) is less than one percentage point greater than for subscribers facing a three-part tariff (90.3%). This relationship is consistent across the entire panel. It is not surprising that downstream content dominates, because popular downstream applications (like video streaming) are bandwidth intensive.

\(^9\)These Gini coefficients are calculated using May 2013 data. For comparison, the Gini coefficient for U.S. income in 2011 is 0.48.
Figure 2.4: *Usage Transition Surface*

**Note:** The level of the surface is the likelihood a subscriber’s usage is in a particular (final) decile of the usage distribution during a particular month, conditional on which (starting) decile that subscriber’s usage was in during the previous month. These statistics reflect usage by subscribers to a single ISP during May 2011 - May 2013. Usage is based upon IPDR data, captured in 15-minute intervals and aggregated to the monthly level.

The three-part tariff is also effective at reducing usage by the heaviest subscribers given the growth rates observed in Table 2.3. While the annualized growth rate for the 99th percentile is 22.8%, usage grows by just 3% between May 2012 and May 2013. This is significantly less than growth in usage by the unlimited subscribers’ (34%). The smaller growth rates at the top of the distribution suggest that three-part tariffs are pushing low-valued content off the network.

The growth rates of the bottom half of three-part tariff subscribers are notably higher than any others. However, these subscribers consume substantially less data per month, so the lower growth rates for the top percentile subscribers still imply much higher yearly increases in GBs because of the higher initial usage. Therefore, these high-usage subscribers are still the primary drivers of network costs.
The differences in growth rates between three-part tariff subscribers and unlimited subscribers are also sizable. There are a couple of explanations for this. First, three-part tariff subscribers enjoy speeds (15.1 Mb/s) more than twice as fast as unlimited subscribers (6.4 Mb/s), which makes it easier to access content. Second, many applications are bandwidth-adaptive. That is, even if the subscriber does nothing behaviorally that would alter their monthly usage, the application will adapt on its own to use whatever bandwidth is available.\footnote{Note also that content providers actively seek to accommodate bandwidth constraints driven by usage allowances. For example, in response to relatively low allowances by some Canadian ISPs, Netflix lowered the default bitrate limit of video from 4800 Kb/s to 625 Kb/s to help users stay under their allowances (OIAC 2013, p. 10).}

Because heavy users generate the majority of network costs, ISPs would ideally generate more revenue from them accordingly. This is possible if usage is persistent. If usage is persistent, only heavy usage users will be affected by the three-part tariff and the amount of cross-subsidization across users will decrease.

To study persistence, I group subscribers by decile of usage for each month. I then create a transition matrix of movement between deciles and plot it in Figure 2.3. The ridge in the Figure shows a subscriber’s current decile is his most likely decile in the subsequent month and that the vast majority of transitions are to a “local” region of the distribution. If a ±2 decile window is used, for example, more than 90% of subscriber transitions are captured.\footnote{Figure 2.3 includes all subscribers, but if the same analysis is performed separately for subscribers on unlimited and three-part tariff plans, the results do not change.} The saddle shape also implies subscribers are least likely to transition to a new decile when they start in a more extreme decile. In fact, I find the tenth decile is the most persistent of all with users remaining there about 65% of the time.

Unsurprisingly, three-part tariffs reduce the amount of subsidization occurring between the bottom and top of the distribution. Table 2.3 shows the distribution of relative prices paid per GB. In May 2013, the average subscriber in the top 10% consumed 15 times more...
data than the average subscriber in the bottom 50%, but paid only 2.3 times more. For unlimited subscribers the results are more extreme. The average subscriber in the top 10% used 21.5 times more and paid 1.1 times more. From this perspective, three-part tariffs are effective at reducing the burden of subsidization on the low-usage subscribers by generating more revenue from high-usage subscribers.

Interestingly, the per-GB cost to a subscriber falls substantially across time. For example, the median user on a three-part tariff plan pays 61.5% less in May 2013 relative to May 2011. This pattern is consistent across all quantiles of the usage distribution for subscribers on both types of plans.

2.4 Intra-Day and Intra-Month Usage

In the three-part tariff plans offered by my ISP, allowances and overages impose implicit prices on users that vary depending upon cumulative usage during that month and on the remaining number of days in the month. Ideally, such prices reduce usage in a way that efficiently saves on network costs. However, allowances that apply to total usage only, such as those used by my ISP, provide no direct incentive by users to reduce peak usage (OIA 2013, p. 16) and may inadvertently reduce off-peak usage as well. Because peak usage alone drives network costs, it is important to know whether such allowances lower peak usage by more than off-peak usage.

To study this issue, I consider usage data for a complete billing cycle during May-June 2012. I start by discussing the data I use to test how three-part tariff allowances affect peak and off-peak usage. I then present direct tests.

Figure 2.4 plots usage for each hour of the day for the average three-part tariff and unlimited subscriber. I report statistics in Megabits per second (Mb/s) to provide perspective on how demanding usage is relative to the maximum potential usage at provisioned speeds.
Figure 2.5: *Average Intra-Day Subscriber Bandwidth Usage*

![Graph showing average bandwidth usage](image)

**Note:** These statistics reflect usage by subscribers to a single ISP during a single billing cycle during May–June 2012. Usage is based upon IPDR data, captured in 15-minute intervals and aggregated to the hourly level.

Just as in the monthly data, the average unlimited subscriber consumes more across the entire day. Notice the pattern of usage observed for both three-part tariff and unlimited subscribers is very similar: early morning hours are the slowest, while the late evening’s are the busiest.

With average usage in Mb/s, Figure 2.4 shows that average usage demands are only a small fraction of a subscriber’s potential (advertised) bandwidth. For example, during the 10 PM hour, the busiest of the day, the average three-part tariff subscriber uses around 0.25 Mb/s, or roughly 0.12 GB per hour. This level of utilization represents around 2% of the maximum possible given the average provisioned connection speed. The low level of utilization is not surprising as many online activities require the subscriber to take time to consume the content even if it arrives very quickly.
Figure 2.6: *Hourly Usage as a Percent of Daily Total*

Note: These statistics reflect usage by subscribers to a single ISP during a single billing cycle during May–June 2012. Usage is based upon IPDR data, captured in 15-minute intervals and aggregated to the monthly level. I denote percentile with “p,” hence the numbers next to p50 denote the subscriber whose usage is at the 50th percentile.
Figures 2.4(a) and 2.4(b) presents the proportion of a subscriber’s daily usage at each hour of the day for subscribers facing a three-part tariff and for unlimited subscribers, respectively. The figures present usage for different quantiles of the distribution (50th, 95th and 99th percentile users). Patterns of usage across the day are nearly identical for three-part tariff and unlimited subscribers. This suggests that those subscribers facing three-part tariffs reduce usage in a proportional manner across the day, since the proportion of daily usage at each hour of the day is the same, but the levels are lower for those facing three-part tariffs as shown in Figure 2.4.

Notably, the heaviest-using subscribers have slightly flatter profiles across the day. This is consistent with such subscribers using more file-sharing applications, like BitTorrent, that generate traffic even if the subscriber is absent. This pattern holds for subscribers on both types (three-part tariff, unlimited) plans.

To test directly for how the three-part affects peak and off-peak usage, I modify and extend the framework used by Nevo et al. (2016) to demonstrate forward-looking behavior by subscribers. Intuitively, subscribers face a dynamic problem within a given billing cycle, regarding how to allocate usage across that billing cycle. A subscriber well below the pace to exceed the usage allowance before the end of the billing cycle has a low probability of exceeding the allowance and consequently a low implicit price of usage. Conversely, a subscriber on pace to exceed the usage allowance has an higher implicit price of usage, but one that is less than the overage price because there is some chance the subscriber’s usage will not reach the allowance. Finally, subscribers that have exceeded the usage allowance at any point in the billing cycle face a constant price, the overage fee.

Hence, when a subscriber’s cumulative usage increases relative to the pace to exceed the usage allowance, then the implicit price of usage increases. If subscribers reduce usage in

\[\text{Notably, my ISP offers a notification system to subscribers regarding the proportion of their allowance that has been consumed up until each point in the billing cycle. Hence, forward-looking behavior is feasible.}\]
a way that lowers the chance of paying overage fees when the probability of overage fees is high, it suggests they are forward-looking and understand how to manage usage across a billing cycle.

In contrast to Nevo et al. (2016), I disaggregate usage into peak (6:00 PM to 11:59 PM) and off-peak hours (otherwise). I then modify the model to capture “peak effects” of implicit prices on usage. If peak effects are non-zero, then subscribers have different sensitivity to implicit prices during these hours. If peak effects are zero, then subscribers respond the same during peak and off-peak periods. If the latter holds, then three-part tariffs do not efficiently target peak usage.

I estimate the following regression model:

$$ln(c_{itp}) = \alpha_i + \alpha_i^{peak}1(p = 1) + [\psi + \psi^{peak}1(p = 1)]x_{it}$$

$$+ \sum_{m=1}^{M=4} \sum_{n=1}^{N=5} [\beta_{nm} + \beta_{nm}^{peak}1(p = 1)]z_{itnm} + \epsilon_{itp}. \tag{2.1}$$

The dependent variable, $ln(c_{itp})$, is the log transformation of subscriber $i$’s peak or off-peak usage at day $t$ in the billing cycle, where $p = 1$ indicates peak usage.$^{13}$ Subscriber fixed effects are denoted by $\alpha_i$ and $\alpha_i^{peak}$; the latter is the “peak” fixed effect. The vector of controls, $x_{it}$, includes day-of-week dummy variables and a billing cycle time trend,$^{14}$ and $\psi^{peak}$ captures the peak effect of controls on usage.

The terms inside the double summation capture a series of interactions that describe a subscriber’s state, such that

$$z_{itnm} = 1[pct_n \leq \bar{c}_{i,t-1} < pct_{n+1}]1[day_m \leq t < day_{m+1}],$$

$^{13}$ Only subscribers for which a complete billing cycle is observed are used in estimation.

$^{14}$ This time trend is identified because subscribers start the billing cycle on different calendar days.
where $\tilde{c}_{i,t-1}$ is the proportion of the usage allowance consumed through $t-1$ days of the billing cycle. The thresholds for the usage allowance indicators are $pct_1 = 0$, $pct_2 = 0.40$, $pct_3 = 0.60$, $pct_4 = 0.80$, $pct_5 = 1$, and $pct_6 = \infty$, while those for the billing day are $day_1 = 10$, $day_2 = 15$, $day_3 = 20$, $day_4 = 25$, and $day_5 = 31$.

I also include a peak indicator here, and interact it with each of the variables. This permits a direct test of whether the response to variation in the possibility of overages is different during peak hours. If $\beta_{nm}^{peak} = 0$ for each ordered pair, $(n, m)$, then I can conclude that subscribers respond similarly to the possibility of overages during each part of the day.

Table 2.4 reports, for each $(n, m)$ ordered pair, the coefficient estimate for the usage allowance interaction in Equation 2.1, $\beta_{nm}$ and for the peak effect, $\beta_{nm}^{peak}$. First, note that none of the peak-effect estimates are statistically or economically significant. This implies that subscribers respond to expected overages by changing behavior uniformly across the day, consistent with the descriptive evidence in Figures 2.4. Thus, the discussion of the impact of three-part tariffs on usage during all hours can focus on the top-half of Table 2.4.

I find a sharp reduction in usage by subscribers with a high probability of exceeding the usage allowance. Subscribers with between 80% and 100% of the usage allowance reduce usage by about 28% over the last six days of the billing cycle. This effect is slightly stronger for those who have already exceeded the usage allowance (33%). Second, subscribers who have consumed only a small proportion of their usage allowance near the end of their billing cycle actually accelerate usage with the assurance that overages are very unlikely. For example, subscribers who have consumed under 40% of the usage allowance increase usage by about 13% over the last six days of the month. Collectively, the results from Table 2.4 show that subscribers are quite sophisticated in their ability to respond to within-month

\[15\text{Note that I interpret these coefficient estimates, due to the log dependent variable, according to Halvorsen and Palmquist (1980). Specifically, } 100 \times (e^{-0.325} - 1) = -27.74\]
variation in the possibility of overages, and that the reduction in usage is quite large on average and similar during peak and off-peak hours.

At the bottom of Table 2.4, I report the proportion of variation explained by the subscriber fixed effect and its interaction with the peak indicator, $\alpha_i$ and $\alpha_{i}^{\text{peak}}$, respectively. Consistent with my earlier discussion, the fixed effects explain about 55% of the total variation in usage, despite the substantial subscriber-specific variation in day-to-day online activities.

By pricing usage similarly at all times of the day, a three-part tariff saves costs by reducing low-value traffic during peak hours but eliminates costless welfare-enhancing usage during off-peak hours. This latter effect harms subscribers and content providers, who lose transactions, without helping the ISP. This suggests that usage-based pricing that specifically incorporates time of day may be more effective at improving overall welfare. The magnitude of any improvement would depend critically on how effectively subscribers can redistribute traffic to off-peak hours (i.e., elasticity across the day). I view this as a crucial topic for future research and strongly encourage experiments that permit clean identification of such elasticities.

\section*{2.5 Conclusion}

I analyze usage data from a North American ISP to study the effects of three-part tariffs, currently the most popular form of usage-based pricing. I find that three-part tariffs are effective at reducing overall usage, generating more revenue from high-usage subscribers and reducing inequality in usage and fees per GB. I also find subscriber usage to be highly persistent, which means usage-based pricing plans such as three-part tariffs can effectively target high-usage subscribers.
Yet, three-part tariffs have minimal effects on the timing of usage. Thus, such plans reduce off-peak usage, causing pure deadweight losses. This suggests that ISPs may be able to further enhance welfare by differentially pricing peak and off-peak usage. Notably, at least one ISP (Exede) has experimented with time-sensitive caps (OIAC 2013, p. 16, footnote 23).\footnote{http://www.dslreports.com/shownews/Exede-Caps-Lifted-For-Overnight-Use-120776} It will be interesting to see whether additional ISPs follow suit.

My results leave many important related questions unanswered. For example, I do not identify granular details of what applications subscribers use. While it seems intuitive that three-part tariffs would tend to remove more high-bandwidth usage like video streaming, further research will be necessary to quantify these effects.

In addition, when ISPs upgrade networks they avoid (or at least mitigate) network congestion, but it is not well known exactly what costs they are avoiding. To quantify completely the welfare effects of reducing network upgrades, it would be useful to estimate costs from congestion. A key challenge is to find data where subscribers face usage prices and congestion simultaneously.

Finally, broadband access is typically bundled with television service from many large cable companies. Given recent moves toward consolidation in the cable industry, such as the recently proposed Charter-Time Warner merger, it appears that cable companies are under pressure to improve their competitive positions. Understanding the welfare implications of such changes, and how they might affect decisions to use usage-based pricing of broadband, are unexplored and important to understand. I look forward to further progress in the area.
Table 2.2: Summary Statistics of Monthly Usage

<table>
<thead>
<tr>
<th></th>
<th>Three-Part Tariff</th>
<th></th>
<th></th>
<th>Annualized Growth Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>May 2011</td>
<td>May 2012</td>
<td>May 2013</td>
<td></td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>21.4</td>
<td>40.1</td>
<td>48.9</td>
<td>0.511</td>
</tr>
<tr>
<td><strong>p25</strong></td>
<td>2.8</td>
<td>7.1</td>
<td>10.7</td>
<td>0.955</td>
</tr>
<tr>
<td><strong>p50</strong></td>
<td>8.9</td>
<td>20.7</td>
<td>29.8</td>
<td>0.830</td>
</tr>
<tr>
<td><strong>p75</strong></td>
<td>25.0</td>
<td>52.4</td>
<td>68.0</td>
<td>0.649</td>
</tr>
<tr>
<td><strong>p90</strong></td>
<td>55.7</td>
<td>103.1</td>
<td>120.7</td>
<td>0.472</td>
</tr>
<tr>
<td><strong>p95</strong></td>
<td>83.9</td>
<td>145.7</td>
<td>160.6</td>
<td>0.384</td>
</tr>
<tr>
<td><strong>p99</strong></td>
<td>164.8</td>
<td>241.1</td>
<td>248.6</td>
<td>0.228</td>
</tr>
<tr>
<td>Usage by Downloading (%)</td>
<td>88.7</td>
<td>89.9</td>
<td>90.3</td>
<td></td>
</tr>
<tr>
<td>Subscribers</td>
<td>48,894</td>
<td>59,550</td>
<td>69,600</td>
<td></td>
</tr>
<tr>
<td><strong>Unlimited</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>32.8</td>
<td>46.6</td>
<td>65.4</td>
<td>0.412</td>
</tr>
<tr>
<td><strong>p25</strong></td>
<td>4.7</td>
<td>7.3</td>
<td>10.1</td>
<td>0.466</td>
</tr>
<tr>
<td><strong>p50</strong></td>
<td>14.6</td>
<td>22.7</td>
<td>33.4</td>
<td>0.513</td>
</tr>
<tr>
<td><strong>p75</strong></td>
<td>39.3</td>
<td>60.6</td>
<td>87.1</td>
<td>0.489</td>
</tr>
<tr>
<td><strong>p90</strong></td>
<td>83.1</td>
<td>119.9</td>
<td>165.2</td>
<td>0.410</td>
</tr>
<tr>
<td><strong>p95</strong></td>
<td>127.0</td>
<td>173.1</td>
<td>230.8</td>
<td>0.348</td>
</tr>
<tr>
<td><strong>p99</strong></td>
<td>243.8</td>
<td>299.0</td>
<td>401.0</td>
<td>0.282</td>
</tr>
<tr>
<td>Usage by Downloading (%)</td>
<td>89.7</td>
<td>90.8</td>
<td>90.9</td>
<td></td>
</tr>
<tr>
<td>Subscribers</td>
<td>28,075</td>
<td>17,426</td>
<td>11,761</td>
<td></td>
</tr>
<tr>
<td><strong>K-S Test Statistic</strong></td>
<td>0.123**</td>
<td>0.037**</td>
<td>0.074**</td>
<td></td>
</tr>
<tr>
<td><strong>Download % DIM t-statistic</strong></td>
<td>11.3**</td>
<td>8.9**</td>
<td>5.4**</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** These statistics reflect usage by subscribers to a single ISP during May 2011 - May 2013. Usage is based upon IPDR data, captured in 15-minute intervals and aggregated to the monthly level. I denote percentile with “p,” hence the numbers next to p25 denote usage by the subscriber whose usage is at the 25th percentile. All usage statistics are in gigabytes. K-S = Kolmogorov-Smirnov and DIM = difference in means. Asterisks denote significance at the 0.05 (*) and 0.01 (**) levels.
Table 2.3: *Summary Statistics of Monthly Revenue*

<table>
<thead>
<tr>
<th></th>
<th>May 2011</th>
<th>May 2012</th>
<th>May 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Three-Part Tariff</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Revenue ($)</td>
<td>83.79</td>
<td>81.44</td>
<td>84.74</td>
</tr>
<tr>
<td>p75 ($/GB)</td>
<td>18.56</td>
<td>8.82</td>
<td>6.21</td>
</tr>
<tr>
<td>p50 ($/GB)</td>
<td>6.39</td>
<td>3.34</td>
<td>2.46</td>
</tr>
<tr>
<td>p25 ($/GB)</td>
<td>4.27</td>
<td>1.57</td>
<td>1.28</td>
</tr>
<tr>
<td>p10 ($/GB)</td>
<td>2.00</td>
<td>1.00</td>
<td>0.86</td>
</tr>
<tr>
<td>p5 ($/GB)</td>
<td>1.33</td>
<td>0.79</td>
<td>0.74</td>
</tr>
<tr>
<td>p1 ($/GB)</td>
<td>0.82</td>
<td>0.65</td>
<td>0.64</td>
</tr>
<tr>
<td>Subscribers</td>
<td>48,894</td>
<td>59,550</td>
<td>69,600</td>
</tr>
<tr>
<td><strong>Unlimited</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Revenue ($)</td>
<td>44.91</td>
<td>44.65</td>
<td>44.39</td>
</tr>
<tr>
<td>p75 ($/GB)</td>
<td>9.21</td>
<td>5.93</td>
<td>4.24</td>
</tr>
<tr>
<td>p50 ($/GB)</td>
<td>3.01</td>
<td>1.92</td>
<td>1.29</td>
</tr>
<tr>
<td>p25 ($/GB)</td>
<td>1.12</td>
<td>0.73</td>
<td>0.49</td>
</tr>
<tr>
<td>p10 ($/GB)</td>
<td>0.53</td>
<td>0.36</td>
<td>0.26</td>
</tr>
<tr>
<td>p5 ($/GB)</td>
<td>0.35</td>
<td>0.26</td>
<td>0.19</td>
</tr>
<tr>
<td>p1 ($/GB)</td>
<td>0.19</td>
<td>0.15</td>
<td>0.11</td>
</tr>
<tr>
<td>Subscribers</td>
<td>28,075</td>
<td>17,426</td>
<td>11,761</td>
</tr>
</tbody>
</table>

**Note:** These statistics reflect usage and expenditures by subscribers to a single ISP during May 2013. Usage is based upon IPDR data, captured in 15-minute intervals and aggregated to the monthly level. Expenditures per GB reflect subscriber plan choices for that month. I denote percentile with “p,” hence the numbers next to p25 denote the subscriber whose $/GB is at the 25th percentile.
Table 2.4: Main Regression Estimates

<table>
<thead>
<tr>
<th>Interactions ($\beta_{nm}$)</th>
<th>$10 \leq t &lt; 15$</th>
<th>$15 \leq t &lt; 20$</th>
<th>$20 \leq t &lt; 25$</th>
<th>$25 \leq t &lt; 31$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0 \leq \tilde{c}_{i,t-1} &lt; 0.40$</td>
<td>-0.041** (0.008)</td>
<td>-0.026** (0.009)</td>
<td>0.039** (0.010)</td>
<td>0.123** (0.010)</td>
</tr>
<tr>
<td>$0.40 \leq \tilde{c}_{i,t-1} &lt; 0.60$</td>
<td>0.002 (0.023)</td>
<td>-0.076** (0.020)</td>
<td>-0.084** (0.018)</td>
<td>-0.001 (0.017)</td>
</tr>
<tr>
<td>$0.60 \leq \tilde{c}_{i,t-1} &lt; 0.80$</td>
<td>-0.008 (0.046)</td>
<td>-0.156** (0.031)</td>
<td>-0.160** (0.024)</td>
<td>-0.080** (0.021)</td>
</tr>
<tr>
<td>$0.80 \leq \tilde{c}_{i,t-1} &lt; 1$</td>
<td>-0.181* (0.078)</td>
<td>-0.257** (0.048)</td>
<td>-0.325** (0.035)</td>
<td>-0.371** (0.028)</td>
</tr>
<tr>
<td>$1 \leq \tilde{c}_{i,t-1}$</td>
<td>-0.079 (0.080)</td>
<td>-0.324** (0.061)</td>
<td>-0.406** (0.042)</td>
<td>-0.460** (0.033)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Peak Effects ($\beta_{peak}$)</th>
<th>$10 \leq t &lt; 15$</th>
<th>$15 \leq t &lt; 20$</th>
<th>$20 \leq t &lt; 25$</th>
<th>$25 \leq t &lt; 31$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0 \leq \tilde{c}_{i,t-1} &lt; 0.40$</td>
<td>-0.015 (0.012)</td>
<td>-0.003 (0.014)</td>
<td>0.023 (0.016)</td>
<td>-0.017 (0.016)</td>
</tr>
<tr>
<td>$0.40 \leq \tilde{c}_{i,t-1} &lt; 0.60$</td>
<td>0.018 (0.039)</td>
<td>-0.025 (0.031)</td>
<td>0.039 (0.029)</td>
<td>0.007 (0.026)</td>
</tr>
<tr>
<td>$0.60 \leq \tilde{c}_{i,t-1} &lt; 0.80$</td>
<td>-0.021 (0.074)</td>
<td>0.055 (0.049)</td>
<td>0.050 (0.038)</td>
<td>-0.031 (0.034)</td>
</tr>
<tr>
<td>$0.80 \leq \tilde{c}_{i,t-1} &lt; 1$</td>
<td>0.006 (0.117)</td>
<td>-0.004 (0.077)</td>
<td>-0.005 (0.055)</td>
<td>-0.007 (0.045)</td>
</tr>
<tr>
<td>$1 \leq \tilde{c}_{i,t-1}$</td>
<td>-0.003 (0.130)</td>
<td>-0.019 (0.095)</td>
<td>-0.001 (0.067)</td>
<td>-0.057 (0.053)</td>
</tr>
</tbody>
</table>

Proportion of Variation Explained by $\alpha_i$ and $\alpha_{\text{peak}}$ 0.550

Note: This table reflects estimates of selected parameters from equation (1). The rows reflect the proportion of a subscriber’s allowance consumed prior to the given period, while the columns reflect the day of the month. I use data from a complete billing cycle during May–June 2012. Usage is based upon IPDR data, captured in 15-minute intervals and aggregated to peak (6:00 PM–11:59 PM) and off-peak (12:00 AM – 5:59 PM) categories. Asterisks denote significance at the 0.05 (*) level and 0.01 (**) level. Robust standard errors are in parentheses.
Chapter 3

The Tragedy of the Last Mile: Congestion Externalities in Broadband Networks

3.1 Introduction

The Internet is now an ever-present part of society, and the demand for online content, especially over-the-top (OTT) video, is soaring. Internet Service Providers (ISPs) choose to invest and meet this demand when there is incentive to do so. An industry estimate places private broadband investment around $1.3 trillion between 1996 and 2013, or about $75 billion per year.\(^1\) Historically broadband investment has been financed by private firms, but its importance is now leading some local governments to pursue municipal broadband and other public funding to support further investment and competition.

In this paper, I provide a key input to understanding how responsive consumers are to network congestion by estimating demand using a novel data set and variation in network congestion and prices. Congested areas of the network are prime candidates for investment because a consistently poor network could lead to welfare losses for consumers, ISPs, and third-parties. In particular, I focus on congestion abatement and its value to consumers. I believe these results are of particular importance to any public policy debate that evaluates the value created by broadband investment. For example, as a part of the Charter/Time Warner Cable merger review, broadband investment to modernize and expand the network is a likely condition for approval.2

The unique data at the center of this work are made available by a North American ISP. These data include hourly observations of Internet usage and network conditions for roughly 45,000 subscribers from February 2015 through December 2015. At the daily level, I am able to uniquely map an account to a cable modem and active Internet plan. For each Internet plan, I observe the price, advertised speeds3 (downstream and upstream), usage allowance, and overage fees. All data tiers charge for data overages at the same per gigabyte (GB) rate. The average subscriber in my data uses 2.3 GB per day, pays $58.89 for a 22 megabit per second (Mbps) downstream connection, and a 267 GB monthly usage allowance.

Network congestion occurs when demand pushes or exceeds the network’s limitations – similar to how video streaming performance degrades if too many people utilize the same WiFi connection. Aside from downstream and upstream speeds, the Federal Communications Commission (FCC) recognizes latency (how long it takes requests to move across the Internet) and packet loss (roughly, the percentage of requests that fail to make it to their

3Speed is measured in Megabits per second (Mbps). For reference, a 20 Mbps downstream connection would download a 4 gigabyte file, or roughly one high-definition movie, in about 27 minutes.
destination) as two important metrics of network performance. Buffering video streams, websites failing to load, and being disconnected from an online video game are common examples of how congestion might affect a consumer. Moreover, due to differences in implementation, such activities as sending email are more resilient to congestion than others like video streaming.

ISP\'s can invest in the physical network in two primary ways. First, network investment can expand the current network, usually in rural and poorer communities, where costs can be greater and broadband is rarer. The government actively promotes this type of investment. Some communities such as Chattanooga, TN and Lafayette, LA have voted for municipal broadband, arguing it promotes competition and investment. Additionally, many politicians have plans similar to President Obama\'s ConnectHome that address how the government will support network investment in these areas.

Second, an ISP can invest in the existing network by increasing capacity and speeds. Here, existing customers are receiving a better experience from the improved network quality. Network performance is of importance to the FCC, which it tracks in its annual Measuring Broadband America reports. In these reports, various network test results of several popular US ISPs are released to promote transparency in the quality and options of broadband available to subscribers. This type of network improvement is what typically abates congestion and is most relevant to this research.

---

5 A third way an ISP could invest in the network is by improving “upstream” relations through various peering and interconnection agreements. Typically, these agreements result in faster, less-congested routes between certain destinations – for example, an ISP and Netflix or Hulu.
7 The program\'s website can be found at http://connecthome.hud.gov and a White House release summarizing the program is found at https://www.whitehouse.gov/the-press-office/2015/07/15/fact-sheet-connecthome-coming-together-ensure-digital-opportunity-all.
A New York Times article\(^9\) describes network investment as having two types of costs: the cost of connecting people’s houses, and the cost of delivering bandwidth to these networks. In general, updating the links between people is the more costly of the two, since a node\(^{10}\) only provides a fixed amount of shared bandwidth to subscribers. Estimating the cost and structure of bandwidth prices are more complicated since they depend on an ISP’s ownership of infrastructure, peering, and interconnection agreements.

*Node splits* are a common way ISPs invest in the core network to improve capacity and lower congestion for a group of subscribers. A node is a common place for bottlenecks to occur and are what commonly demarcate local “last mile” networks. When a node is split, its subscribers are distributed evenly across two new nodes, where network conditions should be improved. Many operators target nodes to be split once average utilization exceeds certain thresholds. I observe five node splits in my data and use these events to compare before- and-after congestion and subscriber usage. After a split, average daily usage increases by 7\% and packet loss, my measure of congestion, drops by 27\%. This suggests there is value to consumers from a less congested network.

The value to consumers of less congestion only captures part of the rent created by the investment: some goes back to the ISP and the rest goes to third-parties. In fact, since the ISP is unable to fully capture these rents, private investment is marginally discouraged. Moreover, recent Title II and net neutrality regulation by the FCC has created uncertainty on the future of the industry, which could depress future investment. Tom Wheeler, the current FCC chairman, declares Title II will have no effect on investment, while other commissioners are doubtful.\(^{11}\)

\(^{10}\)A *node* is a network device that connects a group of subscribers to the rest of the operator’s network.  
My model of subscriber Internet consumption builds on the framework of Nevo et al. (2016) with the notable difference being the inclusion of network congestion and its impact on plan choice and consumption. Similarly, my estimation relies on variation in prices and speeds across plans and (shadow) price variation across the billing cycle that is created by usage-based pricing. I also utilize variation in a subscriber’s observed packet loss to estimate the effect of congestion.

The price variation arising from usage-based pricing is a result of its three-part tariff structure: a subscriber pays a fixed fee each month, and if the associated usage allowance is exceeded, she is charged at a per GB rate thereafter. While overage fees are only assessed if a subscriber exceeds the usage allowance, a forward-looking subscriber understands today’s consumption marginally increases the likelihood of exceeding the usage allowance before the end of the billing cycle – this is a function of how many days remain in the billing cycle and what fraction of the usage allowance has been used previously in the cycle. I incorporate these dynamics in my model similar to Nevo et al. (2016) by allowing consumers to make daily consumption decisions across a billing cycle.

I also use variation in network congestion to identify a subscriber’s sensitivity to poor network states. My hourly data contain packet loss, or the percentage of total packets requested that are either dropped or delayed, at the subscriber level, which I use to proxy for network congestion. As mentioned previously, packet loss is one statistic the FCC uses to benchmark network performance across ISPs.

There are four main advantages to using packet loss over other variables such as a node’s utilization to measure congestion in my model. First, I observe packet loss at a subscriber level, so it is not an aggregate network statistic. Second, I observe wide cross-sectional variation in packet loss across subscribers. Third, packet loss is positively correlated with other common congestion variables. Fourth, my ISP invested in its existing network throughout the year, so I am able to exploit time series variation in my panel as well. My ISP’s net-
work would rate as the third worst in average subscriber packet loss in the latest Measuring Broadband America. Coupled with the aforementioned core network investments by the operator over 2015, this sample offers a wide range of variation in network conditions, ideal for this analysis.

I estimate this finite horizon, dynamic choice model by solving the dynamic problem once for a large number of types. The solution to these dynamic problems is then used to estimate the distribution of types over my sample by minimizing the error between observed and optimal behavior across types. In general, the estimated marginal and joint distributions illustrate the strength of the flexibility built into my estimation approach. Compared to Nevo et al. (2016)’s concentrated type distribution, mine is much more uniform.

These demand estimates are used to measure the value to subscribers when network congestion is eliminated entirely. This is the case where a subscriber’s provisioned speed is always realized. I find the improved network conditions encourage some subscribers to downgrade to cheaper plans, but the loss in revenue from this is entirely offset by an increase in consumer surplus. Subscribers’ realized speeds increased by roughly 19% with each additional Mbps of speed being valued at roughly $2.87. These results suggest that when public policy focuses on network investment, congestion abatement should be considered because of this positive value enjoyed by subscribers.

This paper is most closely related to a literature that studies the demand of residential broadband. Recent examples are Nevo et al. (2016), Malone et al. (2016), and Malone et al. (2014) that use similar high-frequency data to study subscriber behavior. However, this literature dates back to the early 2000s with Varian (2002) and Edell and Varaiya (2002), who run experiments where consumers face different prices for varying allowances and speeds. Goolsbee and Klenow (2006) estimate the benefit to residential broadband; Hitt and Tambe (2007) show Internet usage increases by roughly 22 hours per month when broadband is
introduced. Other related papers are Lambrecht et al. (2007), Dutz et al. (2009), Rosston et al. (2013), and Greenstein and McDevitt (2011).

There is also related research that studies peak-load pricing. Most of these papers focus on industries such as electricity where there is time-of-day variation in the cost of service. Early treatments include Williamson (1966), Carlton (1977), Bailey (1972), Panzar (1976), and Dansby (1978), although the earliest date back to just after the post-war era. In many of these papers, the basic peak load pricing is tweaked and built on by including things such as stochastic demand, government regulation, and various forms of technology. While this literature is quite robust, there is still current work being done. Joskow and Tirole (2006) study the role of metering and transaction costs in monitoring usage in demand, including references to Internet usage metering, Ham et al. (1997) study the role of selection bias in evaluating the effects of peak load pricing, and Puller and West (2013) estimate the efficiency of peak load pricing and the welfare benefits to retail choice. Bergstrom and Mackie-Mason (1991) summarizes the results of moving from uniform to peak-use pricing within an industry and Joskow and Wolfram (2012) summarizes the previous literature and opportunities for future research.

### 3.2 Data

The data for my analysis come from a representative sample of 46,667 North American broadband subscribers. The metropolitan area where the subscribers are drawn have demographic characteristics that are similar to the entire US population, its average income is within 10% of the national average and the demographic composition is similar to the overall US population. Therefore, I expect the insights from my analysis to have external validity in other North American markets. The data include hourly subscriber usage and details of network conditions for February–December 2015.
My data set is constructed from three primary sources. The first source is Internet Protocol Detail Records (IPDR), which report hourly counts of downstream and upstream bytes, packets passed, and packets dropped/delayed by each cable modem.\textsuperscript{12} IPDR also record a cable modem’s node, a device that connects a set of customers to the rest of the operator’s network. The second data source is average hourly utilization by node. The last is billing records by customer, where service plan details (e.g., speed, usage allowance, prices) are included. These data sets are linked by a customer’s anonymized account number, which maps uniquely to a cable modem by day. Using an account number as my unique identifier allows me to follow customers across hardware changes within the sample.

3.2.1 Sample, Internet Plans, and Subscriber Usage

My panel starts on February 1, 2015 and ends on December 31, 2015 and includes 309,307,896 subscriber-day-hour observations. For each observation, I observe downstream/upstream bytes and the total number of packets passed and dropped/delayed. At a daily frequency, I observe each subscriber’s plan and the mapping of consumers to nodes in the network.

The ISP sells Internet access via a menu of plans with more expensive plans including both faster access speeds and larger usage allowances. Overages are charged on a per GB basis after the usage allowance is exceeded. The relationship between monthly usage (GB) and monthly price ($) across the plans is shown in Figure 3.1. The average subscriber pays $58.89 per month for a 22 Mbps downstream connection with a 267 GB usage allowance. The maximum offered speeds and allowances are consistent with those offered in the US, but few subscribers choose them (as I have observed in the data of other ISPs, too). Exceeding the usage allowance is rare in this sample, only 1.6% of subscriber-month observations are over. This rate of overages is notably lower than the approximately 10% rate reported in Nevo et al. (2016), and is largely due to the recent substantial increase in allowances.

\textsuperscript{12}All cable modem hardware identifiers are hashed to preserve anonymity.
Figure 3.1: *Internet Plan Features*

<table>
<thead>
<tr>
<th>Monthly Broadband Price</th>
<th>Monthly Usage (GB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>~1.0</td>
<td>~1.0</td>
</tr>
<tr>
<td>~2.0</td>
<td>~2.0</td>
</tr>
<tr>
<td>~3.5</td>
<td>~3.5</td>
</tr>
<tr>
<td>~7.0</td>
<td>~7.0</td>
</tr>
</tbody>
</table>

*Note:* This figure represents the relationship between monthly usage and price for the ISP’s four Internet plans. Since this ISP has implemented usage-based pricing, there is a set usage allowance for each plan. Once this usage allowance is exceeded, the subscriber is billed on a per GB basis. The overage rate is the same across all four plans. The label that intersects each plan’s line represents the relative differences in speeds.

Subscribers on more expensive Internet plans use more data on average. In Table 3.1, I observe the daily usage distributions of higher tiers dominate those of lower tiers. This suggests differences in usage between tiers is at least partially driven by preferences for larger usage allowances and, possibly, faster speeds. While the differences in subscriber usage may be stark – for example, Tier 4 subscribers use 485% more data on average than Tier 1 subscribers – these extreme subscribers represent only a modest percentage of the subscriber base. Tier Four subscribers only account for 2.5% of the sample. Over 90% of the *subscriber-day* observations are from Tiers 1 and 2.

Hourly average usage in Figure 3.2 follows a cyclical pattern of maximum usage around 9PM and minimum usage around 4AM. This pattern is similar to what is found in Malone et al. (2014) and Nevo et al. (2016) with IPDR data from 2012. Usage during the 9PM
Table 3.1: Daily Usage Distributions by Internet Plan Tier

<table>
<thead>
<tr>
<th></th>
<th>Tier 1</th>
<th>Tier 2</th>
<th>Tier 3</th>
<th>Tier 4</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.4 GB</td>
<td>3.4 GB</td>
<td>5.4 GB</td>
<td>8.2 GB</td>
<td>2.3 GB</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>2.9</td>
<td>5.0</td>
<td>7.3</td>
<td>10.4</td>
<td>4.5</td>
</tr>
<tr>
<td>25th %tile</td>
<td>0.0</td>
<td>0.3</td>
<td>0.6</td>
<td>1.3</td>
<td>0.1</td>
</tr>
<tr>
<td>Median</td>
<td>0.4</td>
<td>1.5</td>
<td>3.1</td>
<td>5.3</td>
<td>0.6</td>
</tr>
<tr>
<td>75th %tile</td>
<td>1.5</td>
<td>4.7</td>
<td>7.6</td>
<td>11.4</td>
<td>2.7</td>
</tr>
<tr>
<td>90th %tile</td>
<td>4.1</td>
<td>9.0</td>
<td>13.6</td>
<td>19.4</td>
<td>6.7</td>
</tr>
<tr>
<td>95th %tile</td>
<td>6.3</td>
<td>12.5</td>
<td>18.5</td>
<td>26.1</td>
<td>10.2</td>
</tr>
<tr>
<td>99th %tile</td>
<td>12.8</td>
<td>22.3</td>
<td>32.0</td>
<td>46.2</td>
<td>20.3</td>
</tr>
<tr>
<td>N</td>
<td>8,539,830</td>
<td>2,910,234</td>
<td>1,117,680</td>
<td>320,085</td>
<td>12,887,829</td>
</tr>
</tbody>
</table>

Note: This table reports daily usage statistics (of the subscriber-day usage distribution) for the four Internet service plans and entire sample.

peak hour is about 0.2 GB, over four times greater than the day’s trough. Throughout this analysis, I will refer to 6PM–11PM as peak hours and the rest of the day as off-peak hours.

About 40% of daily usage occurs during peak hours, as shown in panel (a) of Figure 3.3, with the 9PM hour accounting for just over 8% of a day’s usage. In panel (b), I present the proportion of usage by hour for different deciles of the total usage distribution. The general shape of panel (a) holds for subscribers, regardless of their overall usage level, with the heaviest subscribers (the 10th decile) having only a slightly flatter profile across the day than the others.

Together, Figures 3.2 and 3.3 show a strong and consistent pattern in usage across the day. This pattern suggests ISPs must invest enough in their network to meet demand or the network will become poor and unreliable. One unique feature of my data is I observe numerous periods of excess demand placed on the network that result in congestion. Additionally, I also observe the ISP make substantial investments to increase the capacity and improve the quality of their core network. The behavioral response of subscribers to variation in
Figure 3.2: *Average Usage by Hour and Direction*

![Figure 3.2](image)

*Note:* This figure presents average hourly downstream and upstream usage in gigabytes.

Figure 3.3: *Statistics of Usage as a Percentage of Daily Total*

(a) Hourly Percentages

(b) Hourly Percentages by Decile

*Note:* This figure presents two figures related to how daily usage is proportionally distributed across the day. In panel (a), aggregate hourly percentages are reported for the entire sample. In panel (b), I conditionally report these hourly percentages for deciles 3, 5, 7, and 10. These deciles are calculated using total usage across the entire panel. Each series sums to 100%.
congestion is of primary interest to my analysis. I next discuss measures of congestion, and behavioral responses to congestion-mitigation efforts by the ISP.

### 3.2.2 Network Congestion and Packet Loss

Network congestion occurs when subscriber demand exceeds some capacity constraint on the network. During congested periods, subscribers may find that websites fail to load or online video buffers multiple times. There are two ways to measure congestion in my data. One is through hourly average node utilization. The node being the primary bottleneck in the “last mile” of an ISP’s network. The second being the hourly proportion of packets dropped/delayed, which I, and others, refer to as *packet loss*.

One advantage of hourly packet loss over node utilization is that packet loss is an individual measure instead of an aggregate one. Even when a node is highly utilized, some subscribers may have a normal experience over the hour. Packet loss occurs when data is undeliverable to a subscriber because current network delivery queues are full. Depending on if the data are dropped or delayed, the subscriber’s computer may have to request the data again, further increasing the time of delivery. Packet loss is more likely to occur when nodes are highly utilized, which I observe in my data.\(^\text{13}\) I only observe node utilization at the hourly level, too, which may not be granular enough to accurately reflect a subscriber’s experience within an hour. However, the performance of the network at the instant the subscriber sends and receives packets will be reflected in subscriber-specific hourly packet loss measures.

As part of the FCC’s efforts to monitor the quality of broadband networks, it produces an annual report on the state of broadband networks titled “Measuring Broadband America Fixed Broadband Report”. In the 2015 version, the FCC includes analysis of data from special (SamKnows) modems, which are distributed across numerous ISP networks. In

\(^{13}\)The two measures, hourly utilization and packet loss, have a correlation coefficient equal to 0.164.
Figure 3.4: Industry Statistics on Packet Loss from FCC’s 2015 Report

Note: This Figure is a reproduction from the FCC’s 2015 Measuring Broadband America Fixed Report (https://www.fcc.gov/reports-research/reports/measuring-broadband-america/measuring-broadband-america-2015), which I have modified to include statistics from my ISP’s sample. From the FCC report, it is not clear exactly how their statistics are calculated, but my personal experience with SamKnows data suggests the statistics are the average of hourly packet loss percentages from a specific test conducted by the modem. I calculate the average for “my ISP” according to this methodology. If I aggregate to a daily level, the average is 0.9%.
addition to serving their normal function, these modems conduct hourly tests to measure the performance of the network. One of these tests seeks to measure packet loss by performing an FTP transfer to designated servers located throughout the US. Figure 8 of their report, which I have reproduced in Figure 3.4, presents statistics on packet loss from each ISP’s network. While it is not clear how the values in their Figure 8 are calculated, my experience working with identical data suggests the reported statistics are average hourly packet loss, where the average is taken across subscribers and hours of the day. In my reproduction the FCC figure, I calculate the same statistic for my ISP. This particular measure of network performance would rate my ISP as the third worst across all types of networks in the FCC data (DSL, cable, fiber, and satellite). If you average the overall percentage of daily packets lost, rather than the average of hourly percentages, my ISP would be the worst at around...
Figure 3.6: *Average Hourly Subscriber Packet Loss*

(a) Average Packet Loss

(b) Variation in Packet Loss

**Note:** These figures report statistics of average hourly packet loss. For each subscriber in the sample, I average hourly packet loss across the panel to generate these figures. In panel (a), I report average hourly packet loss for all subscribers. In panel (b), I report the percentage of average packet loss that is over various thresholds. For each subscriber in the sample, I average hourly packet loss across the panel. The percentage of hourly observations over 0.2%, 0.4%, 0.6%, 0.8%, and 1% are shown in the figure.

0.9% of all packets lost. By either definition, my sample represents a great opportunity to study the importance of congestion in broadband networks.

I now provide a more detailed analysis of the measures of congestion I have in my sample. More packets are passed during peak hours when usage is highest, as shown in panel (a) of Figure 3.5. This relationship follows how requests are made online. When a subscriber requests a website, a file to download, or a video to stream on the Internet, packets are sent between the subscriber’s computer and the item’s location. I also observe the highest frequency of dropped/delayed packets during peak hours in panel (b). The reasoning is twofold. First, there are more packets passed during these hours. Second, peak hours are when node utilization is highest.
At the subscriber level of observation, the distribution of hourly packet loss is highly skewed. For example, in panel (a) of Figure 3.6, average packet loss is around 1% at 9PM. However, from panel (b), I find over 85% experience less than 0.2% packet loss. Therefore, the majority of people experience little packet loss over the day, but in some cases, packet loss is very severe. The effects of packet loss on customer experience can be variable, too. For example, when watching a streaming video, 0.5% of packet loss may be acceptable for the video to finish. However, if someone is browsing a website, dropping a single packet could be the difference in a website failing to load correctly. This is important from a modeling standpoint, as I provide a flexible framework to estimate a rich distribution of tastes, which accounts for heterogeneity in the types of content the individual prefers to consume.

Panel (b) of Figure 3.6 better captures the right-tail of the packet loss distribution. In this Figure, I present the percentage of subscribers that are over various packet loss thresholds by hour. Notice in the early morning, when packet loss is lowest, about 3% of subscribers still experience about 1% packet loss on average, compared to the day’s maximum of 10% during peak hours. Interestingly, after 8AM the percentage of subscribers exceeding each threshold remain fairly constant over the remainder of the day.

### 3.2.3 Evolution of Network Quality

Since my data span a ten-month window, I observe changes in the overall quality of the network that, given the correlation between node utilization and packet loss, would improve packet loss and the network state. In panel (a) of Figure 3.7, the weighted average of peak node utilization is plotted for each week in my panel. Not only is there variation across the year, but there are distinct drops in May, September, and December where the ISP improved network capacity. These changes are also noticeable in how median peak utilization varies in panel (b). The dashed whiskers in panel (b) represent the 5th and 95th percentiles of peak usage, where even during these network events the variation within a week is unaffected.
Figure 3.7: **Weekly Node Utilization Statistics**

(a) Weighted Average Utilization

(b) Variation in Peak Utilization

*Note:* Panel (a) plots the weighted average of peak utilization by week. The weights in this case are the number of people on the node. Panel (b) plots the weekly variation in peak utilization. The green box is the IQR, the red line is the median, and the blue dashed lines extend to the 5th and 95th percentiles.

One way an ISP can alleviate congestion on a node is to perform a *node split*. This is just one option available to an ISP – an ISP can use other hardware, software, and licensing methods to change the capacity of and bandwidth made available to a node. An example of a node split is for the operator to take a node and split its subscribers across two new nodes. When such a change is made, the network state for the affected subscribers should be improved since there are half as many subscribers using the same node. If subscriber behavior is responsive to such changes in network quality, I would expect an increase in usage. Note that the increase in usage could come from a change in the subscriber himself, or bandwidth adaptive applications becoming more responsive.

There are 5 distinct node splits in the data, whereby a group of subscribers is clearly split over two new nodes. Changes in network conditions are summarized in Table 3.2 and subscriber usage in Table 3.3. I do see improvements in the average network state with
Table 3.2: Changes in Node Utilization and Packet Loss After Node Split

<table>
<thead>
<tr>
<th></th>
<th>Before</th>
<th>After</th>
<th>Diff</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hourly Utilization</td>
<td>49%</td>
<td>34%</td>
<td>-15%</td>
<td>-31%</td>
</tr>
<tr>
<td>Max Hourly Utilization</td>
<td>87%</td>
<td>62%</td>
<td>-25%</td>
<td>-29%</td>
</tr>
<tr>
<td>Hourly Packet Loss</td>
<td>0.11%</td>
<td>0.08%</td>
<td>-0.03%</td>
<td>-27%</td>
</tr>
<tr>
<td>Max Hourly Packet Loss</td>
<td>1.0%</td>
<td>0.61%</td>
<td>-0.39%</td>
<td>-39%</td>
</tr>
</tbody>
</table>

Note: This table reports how the averages of node utilization and packet loss compare before and after the node split. 7 days of data is taken from before and after the node split date to calculate means. These averages are at the node level of observation and are weighted by the number of people on the node.

Table 3.3: Changes in Daily Usage After Node Split

<table>
<thead>
<tr>
<th></th>
<th>Before</th>
<th>After</th>
<th>Diff</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Off-Peak Usage</td>
<td>1.51 GB</td>
<td>1.57 GB</td>
<td>0.06 GB</td>
<td>4.0%</td>
</tr>
<tr>
<td>Peak Usage</td>
<td>1.04 GB</td>
<td>1.16 GB</td>
<td>0.12 GB</td>
<td>12.0%</td>
</tr>
<tr>
<td>Total Daily Usage</td>
<td>2.55 GB</td>
<td>2.73 GB</td>
<td>0.18 GB</td>
<td>7.1%</td>
</tr>
</tbody>
</table>

Note: This table reports how subscriber behavior changed around a node split. 7 days of daily usage is taken from before and after the node split date to calculate means. This table summarizes usage for 2,627 subscribers over 5 node splits.

decreases in both utilization and packet loss. Maximum hourly node utilization falls by 29% and maximum hourly packet loss falls by 39%. Over this same period, I find a 7.1% increase in daily usage. Peak usage increases relatively more (12%) than off-peak usage (4.0%). This suggests that there is some degree of unmet demand prior to the node split that is now able to be realized.

In Table 3.4 and Figure 3.8, packet loss is split into seven bins that are used to study how persistent packet loss is day-to-day; the values in the heat map are the same as in the table. From these transition probabilities, there are a couple of notable takeaways. First, if a subscriber’s peak packet loss is poor one day, there is a high probability it will be better the next day. Second, if a subscriber does end up in the worst packet loss state, they are most
Figure 3.8: Heatmap of Peak Packet Loss Transitions

Note: This heatmap graphically reports probabilities of peak hour-day packet loss transitions at the subscriber level of observation. Each bin is of the form \((x\%, y\%)\) and represents a range of packet loss. The first bin includes 0% packet loss, too. The values of each cell are reported in Table 3.4.

Table 3.4: Transition Matrix of Peak Packet Loss

\[
\begin{array}{cccccccc}
\text{Initial State} & 0-0.2 & 0.2-0.4 & 0.4-0.6 & 0.6-0.8 & 0.8-1 & 1-10 & 10-100 \\
0-0.2 & 0.984 & 0.002 & 0.001 & 0.001 & 0.001 & 0.006 & 0.004 \\
0.2-0.4 & 0.662 & 0.086 & 0.044 & 0.027 & 0.021 & 0.124 & 0.037 \\
0.4-0.6 & 0.570 & 0.074 & 0.055 & 0.037 & 0.027 & 0.186 & 0.051 \\
0.6-0.8 & 0.526 & 0.041 & 0.031 & 0.062 & 0.039 & 0.235 & 0.066 \\
0.8-1 & 0.511 & 0.032 & 0.026 & 0.042 & 0.059 & 0.244 & 0.087 \\
1-10 & 0.316 & 0.023 & 0.020 & 0.029 & 0.029 & 0.364 & 0.218 \\
10-100 & 0.122 & 0.004 & 0.003 & 0.005 & 0.005 & 0.119 & 0.741 \\
\end{array}
\]

Note: This table reports probabilities of peak hour-day packet loss transitions at the subscriber level of observation. Each bin is of the form \((x\%, y\%)\) and represents a range of packet loss. The first bin includes 0% packet loss, too.
likely to be in a poor state the next day. Third, the vast majority of subscribers experience low packet loss and will experience low packet loss tomorrow.

For the model, I use the transition matrix in Table 3.4 to estimate the frequencies of transition between packet loss, or network congestion, states. Below in the model discussion, this will be $G_\psi$. This matrix will be used to solve the model. For the estimation procedure, all I need are day-hour observations of daily consumption and the observed peak packet loss state for each account in the sample.

### 3.3 Model

My model builds on the model of Nevo et al. (2016). Like Nevo et al. (2016), I assume a finite horizon, that a subscriber’s discount rate is $\beta$, and that a subscriber makes a consumption decision each period on his optimally chosen plan. The primary difference between the two models is I include network congestion and allow for it to impact subscribers’ plan and consumption choices.

Given that my focus is on the role of congestion, I limit my sample to only subscribers who never switched plans over the duration of the panel. This does not affect my analysis for two reasons. First, service plans were upgraded shortly before my sample, and, second, about 90% of subscribers made no changes during my period of observation. Allowing for plan switching introduces a dependency across billing cycles in the dynamic problem, and by not modeling it the computational burden of solving the model is reduced.

#### 3.3.1 Subscriber Utility From Content

Subscribers derive utility from consumption of content. Each day of a billing cycle, $t = 1,...,T$, a subscriber chooses the amount of content to consume on their chosen service plan, $k = 1,...,K$. Plans are characterized by a provisioned speed content is delivered, $s_k$, by a
usage allowance, \( C_k \), by a fixed fee \( F_k \) that pays for all usage up to the allowance and by an overage price, \( p_k \), per GB of usage in excess of the allowance. The menu of plans, and the characteristics of each, are fixed.\(^{14}\) The provisioned speed is impacted by the state of the network, \( \psi \), which changes daily due to variation in congestion and frequent network upgrades. I assume this evolution follows a first-order Markov process, \( G_\psi \). Estimates of this process are presented in Table 3.4.

Utility from content is additively separable over all days in the billing cycle, and across billing cycles.\(^{15}\) Let daily consumption of content be denoted by \( c \). The utility for a subscriber of type \( h \) on plan \( k \) is given by

\[
  u_{hk}(c, \psi, \upsilon) = \upsilon \left( \frac{c^{1-\alpha_h}}{1 - \alpha_h} \right) - c \left( \frac{\kappa_h}{\ln(\psi s_k)} \right).
\]

The first term captures the subscriber’s utility from consuming the content. Marginal utility is declining, as I expect the first of any activity (email, web browsing, video, etc.) to bring higher marginal utility than subsequent usage. The convexity of the utility function is also quite flexible, nesting everything between log (\( \alpha_h \to 1 \)) and linear (\( \alpha_h = 0 \)). This leads to a straightforward link between \( \alpha_h \) and the price elasticity of demand, such that \( \alpha_h \) is the elasticity with respect to the entire cost associated with consuming content, both monetary and non-monetary. Uncertainty in utility from consumption of content is introduced by a time-varying shock, \( \upsilon \), which is realized on the day the consumption decision is made. I assume that \( \upsilon \) is independently and identically distributed according to a log normal distribution with parameters, \( \mu_h^\upsilon \) and \( \sigma_h^\upsilon \), for each type, \( h \).

\(^{14}\) Plans were changed months prior to my sample, but unchanged during my sample, and the ISP had no plans to change them in the months after my sample ends.

\(^{15}\) In this way, I assume content with a similar marginal utility is generated each day or constantly refreshed. This may not be the case for a subscriber who has not previously had access to the Internet. Below I will assume decreasing marginal utility within a time period, but additive across periods.
The second term captures the subscriber’s non-monetary cost of consuming content. This cost, $\frac{\kappa_h}{\psi \ln(s_k)}$, is time-varying and subscriber-specific. The $\kappa_h > 0$ parameter captures both a subscriber’s preference for speed and the waiting cost of transferring content, which depends on the plan’s provisioned speed and the state of the network. Importantly, for any finite speed, this specification implies that each subscriber type has a satiation point even in the absence of overage charges. Thus, my specification of this cost departs slightly from that of Nevo et al. (2016), by not including an additive fixed value ($\kappa_1$ in their model) that interacts with speed. This parameter is only weakly identified, as the limiting case with unbounded speed that would fully reveal this cost does not occur in my data. The interaction of the network state and speed captures the way in which this operator has chosen to ration bandwidth during congestion, as a proportional degradation of the provisioned speed.\(^\text{16}\)

### 3.3.2 Optimal Usage

The observability of the network state and upgrade of plan features prior to my sample, which limits plan switching and permits focusing my analysis on the approximately 90% of consumers enrolled on a single plan the entire sample period, simplifies the characterization of optimal usage. Specifically, like Nevo et al. (2016), each consumer must solve a finite-horizon dynamic programming problem within each billing cycle. For a subscriber on plan $k$, I denote the amount of his unused usage allowance, on day $t$ of the billing cycle, as $\overline{C}_{kt} \equiv \text{Max}\{\overline{C}_k - C_{t-1}, 0\}$, where $C_{t-1}$ is cumulative usage up until day $t$. Similarly, denote day-$t$ overage as $O_{tk}(c_t) \equiv \text{Max}\{c_t - \overline{C}_{kt}, 0\}$.

In the last day of the billing cycle ($T$), the subscriber faces no intertemporal tradeoffs and solves a static optimization problem, conditional on his cumulative usage $C_{T-1}$ and the realization of the preference shock, $\nu_T$. Once $\nu_T$ is realized, subscribers who will not incur

\(^{16}\)My discussions with network engineers suggest this rationing rule can easily be altered to ration capacity differently during times of congestion.
overage charges (i.e., \( C_{kT} \) is high) consume such that \( \frac{\partial u_h(c_t, y_t, \psi_t; k)}{\partial c_t} = 0 \). If \( \frac{\partial u_h(c_t, y_t, \psi_t; k)}{\partial c_t} \) at \( c_t = C_{kT} \) is positive and less than \( p_k \), then consuming the remaining allowance is optimal. For those subscribers above the allowance (i.e., \( C_{kT} = 0 \) and a high realization of \( \nu_T \)), it is optimal to consume such that \( \frac{\partial u_h(c_t, y_t, \psi_t; k)}{\partial c_t} = p_k \). Denote this optimal level of consumption in each scenario by \( c^*_h(T-1, \nu_T) \). Given this optimal policy for consumption, utility in the terminal period is

\[
V_{hT}(C_{T-1}, \psi_T, \nu_T) = \nu_T \left( \frac{(c^*_h)^{1-\alpha_h}}{1-\alpha_h} \right) - c^*_h \left( \frac{\kappa_h}{\ln(\psi_T s_k)} \right) - p_k O_{tk}(c^*_h) .
\]

For other days in the billing period, \( t < T \), consumption increases cumulative consumption and alters the state, so the optimal policy for a subscriber must incorporate this. The optimal policy for any \( t < T \) can be expressed recursively such that for type \( h \) on plan \( k \)

\[
c^*_h(C_{t-1}, \psi_t, \nu_t) = \arg\max_{c_t} \left\{ u_t \left( \frac{c_t^{1-\alpha_h}}{1-\alpha_h} \right) - c_t \left( \frac{\kappa_h}{\ln(\psi_t s_k)} \right) - p_k O_{tk}(c_t) + E_{\psi} \left[ V_{ht+1}(C_{t+1} + c_t, \psi_{t+1}) \right] \right\} .
\]

Alternatively, defining the shadow price of consumption as

\[
\tilde{p}_h(c_t, C_{t-1}, \psi_t) = \begin{cases} 
p_k & \text{if } O_{tk}(c_t) > 0 \\
\frac{dE_{\psi}[V_{hh+1}(C_{t+1} + c_t, \psi_{t+1})]}{dc_t} & \text{if } O_{tk}(c_t) = 0 .
\end{cases}
\]

the optimal consumption choice in period \( t \) satisfies

\[
c^*_h = \left( \frac{\kappa_h}{\ln(\psi_t s_k)} + \tilde{p}_h(c^*_h, C_{t-1}, \psi_t) \right) \frac{1}{\alpha_h} . \tag{3.1}
\]

The relationship between \( \alpha_h \) and the price elasticity of usage is clear from Equation 3.1. A type with parameter \( \alpha_h \) has a usage elasticity equal to \( -\frac{1}{\alpha_h} \) with respect to the entire marginal cost of content, \( \frac{\kappa_h}{\ln(\psi_t s_k)} + \tilde{p}_h(c_t, C_{t-1}, \psi_t) . \)
The value function associated with the optimal usage policy is

\[ V_{hkt}(C_{t-1}, \psi_t, v_t) = v_t \left( \frac{(c^*_{hkt})^{1-h}}{1-h} \right) - c^*_{hkt} \left( \frac{\kappa_h}{m(\psi_t s_k)} \right) - p_k O_t (c^*_{hkt}) + E_\psi [V_{hkt}(t+1)(C_{t-1} + c^*_{hkt}, \psi_{t+1})] \]

for each 3-tuple, \((C_{t-1}, \psi_t, v_t)\). Then for all \(t < T\), the expected continuation value is

\[ E_\psi [V_{hkt}(C_{t-1}, \psi_{t+1})] = \int_{\psi} \left( \int_{v_t} V_{hkt}(C_{t-1}, \psi_{t+1}, v_t) dG^h_v(v_t) \right) dG^h_\psi(\psi_{t+1}|\psi_t), \]

and the mean of a subscriber’s usage at each observable state is

\[ c^*_{hkt}(C_{t-1}, \psi_t) = \int_{\psi} c^*_{hkt}(C_{t-1}, \psi_t, v_t) dG^h_v(v_t). \quad (3.2) \]

The Markov process associated with the solution to the dynamic program also implies a distribution for the time spent in each state, \((t, C_{t-1}, \psi_t)\), over a billing cycle, \(P_{hk^*_h m}(C_{m-1}, \psi_m)\). This process along with expected consumption at each state form the basis of my estimation algorithm.

### 3.3.3 Plan Choice

I assume subscribers select plans to maximize expected utility, before observing any utility shocks, and remain on that plan during my sample. More precisely, I assume that the subscriber selects one of the offered plans, \(k \in \{1, ..., K\}\), or no plan, \(k = 0\), such that

\[ k^*_h = \arg\max_{k \in \{0,1, ..., K\}} \left\{ E[V_{hk1}(C_1 = 0, \psi_1)] - F_k \right\}. \]

The optimal plan, \(k^*_h\), maximizes expected utility for the subscriber given the current state of the network and optimal usage decisions, \(E[V_{hk1}(C_1 = 0, \psi_1)]\), net of the plan’s fixed
access fee, $F_k$. The outside option is normalized to have a utility of zero. Note, that I assume that there is no error, so consumers choose the plan that is optimal. Similar to Nevo et al. (2016), (admittedly weak) tests of optimal plan choice reveal that it is rare to observe a subscriber whose usage decisions are such that switching to an alternative plan would yield lower total costs at no slower speeds. The weakness of this optimality test is due to the positive correlation between speed and usage allowances of the offered plans (see Figure 3.1). My assumptions on plan choice are easily relaxed in theory, but introduce a substantial additional computational burden. Given the infrequency of both clear ex-post mistakes in choosing a plan and switching of plans, I believe this is a reasonable assumption for my sample.

### 3.4 Estimation

My estimation approach is a panel-data modification of Fox et al. (2011), proposed in Nevo et al. (2016), which I refer to as fixed-grid fixed-effect (FGFE) least squares. The approach exploits the richness of panel data to build upon the fixed-grid random-effects (FGRE) approach of Fox et al. (2011) used by Nevo et al. (2016). In contrast to the FGRE approaches, my FGFE approach permits identification of each subscriber’s type, rather than just the distribution of types, and also allows consideration of moments from the model that are not non-linear in the type-specific population weights. This is advantageous for identification of the model and consideration of richer counterfactual exercises where knowledge of an individual’s type is useful rather than just the distribution of types.
3.4.1 Econometric Objective Function

For each individual, \( i = 1, \ldots, I \), I have a time series of data, \( m = 1, \ldots, M \), which captures usage at a daily frequency on an optimally chosen plan.\(^{17}\) Thus, I have a daily time series for usage, \((c_{i1}, c_{i2}, \ldots, c_{iM})\), for individual \( i \), as well as the accompanying observable portion of the state, \((t_m, C_{m-1}, \psi_m)\) for each \( m = 1, \ldots, M \). From the solution to the model, for each type \( h \), I store two moments associated with usage on the optimally chosen plan, \( c^*_{hk} t(C_{t-1}, \psi_t) \) and \( c^*_{hk} t^2(C_{t-1}, \psi_t) \): expected usage and the expectation of the square of usage at every observable state. Additionally, I calculate the probability of observing a type in a particular state, \( P_{hk} t(C_{t-1}, \psi_t) \).

The goal of the estimation algorithm is to identify which of the \( H \) types’ behavior from the model best match the behavior of each individual, \( i \), over the panel of data. I use a least-squares criteria to compare fit, such that the type \( h \) that best matches to consumer \( i \) is given by

\[
\hat{h}_i = \min_{\{h = 1, \ldots, H\}} \left[ \sum_{m=1}^{M} \tilde{z}_{ih}' \tilde{z}_{ih} \right],
\]

where

\[
\tilde{z}_{ih} = \begin{pmatrix}
    c_{im} - c^*_{hk} t(C_{m-1}, \psi_m) \\
    c^*_{im} - c^*_{hk} t^2(C_{m-1}, \psi_m) \\
    1 - P_{hk} t(C_{m-1}, \psi_m)
\end{pmatrix}.
\]

This process is repeated for each \( i \). Aggregating across the chosen types for each consumer, \( i \), the population weights for each type, \( h \), is then

\(^{17}\)I drop the small fraction of subscribers, less than 2%, for which I do not observe a complete time series.
\[
\hat{\theta}_h = \frac{1}{I} \sum_{i=1}^{I} \mathbf{1} \left[ \hat{h}_i = h \right].
\]

There are a numerous advantages of the FGFE approach in panel-data applications like mine. First, even compared to the constrained convex optimization problem in the FGRE approach of Fox et al. (2011), there can be computational advantages introduced by only searching over a fixed grid of types. Second, the FGFE approach does not require that the moments used in estimation be linear in the type-specific weights. In the FGRE approaches, this linearity is necessary or the problem becomes a constrained nonlinear optimization problem that is intractable with even a moderate number of types. These richer moments can be particularly helpful in identification, as Nevo and Williams (2016) show.

Another advantage of the FGFE approach is that it fully characterizes the discrete distribution of types, but in contrast to FGRE demand models like Fox et al. (2011) and Nevo et al. (2016), the mapping between an individual and a type is preserved (\(\hat{h}_i\)). The panel data eliminates the need to aggregate across consumers to form moments, and this permits an individual’s type to be inferred, rather than just the distribution of types in the population. In many applications in Industrial Organization, inferring an individual’s type rather than only the distribution of types is useful. From the firms’ perspective, this may permit different forms of discrimination (third-degree rather than second-degree) to be implemented through targeted offerings. Knowledge of each individual’s type can also permit a decomposition of the parameters via the minimum-distance procedure of Chamberlain (1982) when observable characteristics of the individual are available, as is done in Nevo (2001). For example, one can regress the parameters describing an individual’s type, \((\mu_{\hat{h}_i}^\nu, \sigma_{\hat{h}_i}^\nu, \alpha_{\hat{h}_i}, \kappa_{\hat{h}_i})\), on a vector of the individual’s observed characteristics to decompose the parameter into observable and un-
observable determinants of preferences. This may be particularly useful in labor and health applications where a rich set of observable characteristics are often available.

3.4.2 Identification

Identification of my model closely follows the discussion in Nevo et al. (2016). There are a few important differences, each simplifying and improving identification. First, I eliminate the $\kappa_{1h}$ parameter that led to a satiation point for usage even when speed was unbounded. While I observe higher speeds than Nevo et al. (2016), this dimension to the type space is not needed because the limiting case is clearly not in my data, and given the additional computational complexity of my model, eliminating it permits me to consider a denser grid over the other parameters. Second, like Nevo et al. (2016), usage and plan choices are strong sources of identification. Plan choice can be thought of as assigning each type to a plan and putting a uniform prior over the types on each plan, while the usage moments can then distinguish between the types choosing a plan. The flexibility of the FGFE approach is also important here, as I am able to consider richer usage moments, because I am not restricted to moments that preserve linearity in the type weights. This is how I am able to consider the first and second conditional moments of usage, in contrast to Nevo et al. (2016), which only uses the unconditional first moment. Finally, I also have an additional source of variation in the price of usage that can help identify a type. Like Nevo et al. (2016), usage-based pricing is particularly helpful, as I observe a large number of marginal decisions by each consumer, weighing the benefit of consuming more content against the increase in the probability of overages (i.e., the shadow price of usage). This variation is helpful for pinning down the primary determinant of an individual’s elasticity of demand, $\alpha_{\hat{h}_i}$. However, in addition to this price variation introduced by the nonlinear pricing, I also have extensive variation in the network state, which shifts the cost of consuming content. This is helpful in pinning down
an individual’s preference for speed, $\kappa_{\hat{h}_i}$, which would otherwise largely be identified by plan choice alone.

3.5 Results

I present my estimation results in two parts. First, I report my estimates of the types distributions. Next, I discuss the results of a counterfactual exercise that measures the value to consumers of eliminating all congestion on the network.

3.5.1 Type Distribution Estimates

I estimate a weight greater than 0.01% for 164 types. That is, 164 different types $h$ were chosen for at least one subscriber ($i$), or 164 different $\hat{h}_i$ were chosen among all possible types. Conditional on being chosen, I find the weights are distributed rather uniformly among the types. This is in contrast to Nevo et al. (2016), which finds a concentrated distribution of types. Their most common type accounted for 28% of the total mass, the top five types accounted for 65%, the top 10 for 78% and the top 20 for 90%. Nevo and Williams (2016) show this difference in the concentration of the types tend to be largely due to the difference between the FGFE employed here and the FGRE effects approach employed by Nevo et al. (2016).

Interestingly, the much larger number of types I estimate here is almost exclusively due to the most expensive plan, which accounts for 110 of the 164 positive types. On the most expensive plan, there is a wide variety of behavior that must be explained. I observe many low usage subscribers, which when optimal plan choice is assumed, can only be rationalized by a type with an intense preference for speed – think of an individual that video conferences or occasionally downloads very large files, and wants the applications used to perform
Figure 3.9: Estimated Marginal Distributions of Model Parameters

(a) $\mu$

(b) $\sigma$

(c) $\alpha$

(d) $\kappa$

Note: The figures are the estimated marginal distributions of the four parameters in my model. Within each panel, the sum of all five bars in each figure total 100%.
Table 3.5: Descriptive Statistics for Types

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_h$</td>
<td>0.759</td>
<td>0.75</td>
<td>1</td>
</tr>
<tr>
<td>$\sigma_h$</td>
<td>0.610</td>
<td>0.6</td>
<td>0.9</td>
</tr>
<tr>
<td>$\kappa_h$</td>
<td>15.709</td>
<td>15.5</td>
<td>15.5</td>
</tr>
<tr>
<td>$\alpha_h$</td>
<td>0.483</td>
<td>0.5</td>
<td>0.375</td>
</tr>
</tbody>
</table>

Note: This table reports descriptive statistics of the type distribution: mean, median, and mode.

seamlessly. Similarly, I have many individuals that desire a large allowance, but have a less-intense preference for speed.

Figure 3.9 presents the marginal distributions for each of the parameters, $(\mu_h, \sigma_h, \kappa_h, \alpha_h)$. Interestingly, the type distribution for each parameter is quite uniform, or non-normal, across the support. This contrasts the lumpy distributions recovered by Nevo et al. (2016). Therefore, the mean, median, and mode of the four parameters, which are reported in Table 3.5, are quite similar. For each of the parameters, the mean, median, and mode are within 10% of one another.

The estimated joint distributions are much more irregular, neither uniform or normal. These joint distributions for each combination of the parameters, six in total, are presented in Figure 3.10. Like Nevo et al. (2016), the joint distributions are multi-peaked and vary considerably by the pair of parameters considered. This demonstrates the importance of the flexibility of my estimation approach, which allows for free correlations between each pair of parameters rather than the zero covariance often assumed in structural econometric applications. This flexibility is reflected in the fit of the model. For all plans, the correlation between the empirical moments and the fitted moments is above 90%. The model also fits patterns in the data not explicitly used in estimation, similar to those reported in Nevo et al. (2016).
Figure 3.10: Estimated Joint Distributions of Model Parameters

Note: The figures are the estimated joint distributions of the four parameters in my model. Within each panel, the sum of all twenty-five bars in each figure total 100%.
Table 3.6: Counterfactual Results from Eliminating Network Congestion

<table>
<thead>
<tr>
<th>Usage and Surplus</th>
<th>Current Offerings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>With Congestion</td>
</tr>
<tr>
<td>Daily Usage (GB)</td>
<td>2.2 GB</td>
</tr>
<tr>
<td>Provisioned Speed (Mbps)</td>
<td>22.3</td>
</tr>
<tr>
<td>Realized Avg. Speed (Mbps)</td>
<td>16.7</td>
</tr>
<tr>
<td>Consumer Surplus ($)</td>
<td>65.80</td>
</tr>
<tr>
<td>Revenue ($)</td>
<td>58.58</td>
</tr>
</tbody>
</table>

Note: This table reports results from the counterfactual exercise of eliminating network congestion, i.e., provisioned speeds are always delivered to subscribers.

3.5.2 Value of Eliminating Congestion

In this counterfactual exercise, I measure the value to subscribers from eliminating network congestion, or the case where provisioned speeds are always realized to subscribers. This exercise is important because it illustrates the value, or lack thereof, of many types of core network improvements that go towards reducing congestion and better meeting demand.

By eliminating congestion, I estimate an increase in consumer surplus of 14%, as shown in Table 3.6. Daily usage increases from 2.2 GB/day to 2.5 GB/day, or about an additional 9 GB over a 30-day billing cycle. I also observe many subscribers downgrading plans as a result of the improved network state. Since the speeds of a cheaper plan are now guaranteed, subscribers with a stronger preference for speed over usage may be better served on the cheaper plan. This movement to cheaper plans lowers average revenue for the ISP, but the increase in consumer surplus is large enough to offset this drop: $4.24 drop in revenue and $8.90 increase in consumer surplus for a net difference of $4.66. Since subscribers are receiving speeds roughly 19% faster, I estimate a subscriber values each additional Mbps of realized speed at $2.87.
3.6 Conclusion

I estimate demand for residential broadband using a ten-month panel of hourly subscriber usage and network conditions. The key feature of my model is the incorporation of network congestion and allowing it to affect a subscriber’s daily consumption decision.\textsuperscript{18} There are three sources of variation I exploit in my data. First, I use (shadow) price variation that results from the structure of usage-based pricing’s three-part tariff. Second, I use cross-sectional variation in packet loss, my measure of network congestion, across subscribers. Third, my ISP invested in the core network several times throughout 2015, creating times series variation in the overall quality of the network.

My demand estimates are used to measure the value to subscribers from eliminating network congestion. I find the improved network conditions encourage some subscribers to downgrade, but any loss in revenue is entirely offset by an increase in consumer surplus. Subscribers’ realized speeds increased by roughly 18\% with each additional Mbps of speed being valued at roughly $2.87.

There are several extensions to the basic model in this paper that can be explored in future versions. First, I could allow consumers to switch plans. In my sample, around 12\% of subscribers make a plan change by either moving to a more expensive or cheaper plan; these subscribers are omitted from my original estimation. Under a usage-based pricing regime, I expect some consumers may upgrade to a plan with a larger usage allowance to account for their growing demand, while others may downgrade to better align their usage with a lower allowance or due to cost concerns.

Next, I could allow consumers to make two consumption choices each day: one for off-peak hours (12AM–5PM) and peak hours (6PM–11PM). This permits two additional pieces of analysis. First, I am able to study how congestion differentially affects usage during different

\textsuperscript{18}This differs from previous research such as Nevo et al. (2016).
times of the day. Since demand is greater during peak hours, I may expect subscribers to behave differently. I am also able to explore the welfare implications of peak-use pricing as an implementation of usage-based pricing. Large demand during peak hours drives higher service costs for the ISP during these hours. However, during off-peak hours, service is mostly costless since the network is less congested. For example, if only peak hour usage counted towards the usage allowance, I may observe subscribers flattening their usage profile across the day by shifting peak hour usage to off-peak times.

Then, in related research, I believe there are many related topics. First, my hourly data usage is aggregated across all traffic types. Getting high-frequency data that are disaggregated by application or traffic type would permit a more detailed analysis and understanding of how congestion differentially affects each type. Moreover, moving to a more granular level, say 5 to 15 minutes, would allow for a more exact understanding of the correlation between usage and congestion. Second, residential broadband and traditional linear television (TV) services are closely related and are often bundled together by ISPs. In future research, I hope to obtain linear TV data in conjunction with disaggregated high-frequency Internet usage by type to more completely explore how subscribers use broadband Internet. Specific relationships like the substitutability of linear TV and over-the-top video (OTTV) and how linear TV usage and network congestion are correlated could be explored with such a data set.

\[19\] This type of usage-based pricing was common with the telephone companies when they offered unlimited minutes on nights and weekends.
Bibliography


