Chapter 16

The Empirics of the Digital Divide: Can Duration Analysis Help?

Wei-Min Hu
Peking University HSBC School of Business, China

James E. Prieger*
Pepperdine University, USA

ABSTRACT

Accurate measurement of digital divides is important for policy purposes. Empirical studies of broadband subscription gaps have largely used cross-sectional data, which cannot speak to the timing of technological adoption. Yet, the dynamics of a digital divide are important and deserve study. With the goal of improving our understanding of appropriate techniques for analyzing digital divides, we review econometric methodology and propose the use of duration analysis. We compare the performance of alternative estimation methods using a large dataset on DSL subscription in the U.S., paying particular attention to whether women, blacks, and Hispanics catch up to others in the broadband adoption race. We conclude that duration analysis best captures the dynamics of the broadband gaps and is a useful addition to the analytic tool box of digital divide researchers. Our results support the official collection of broadband statistics in panel form, where the same households are followed over time.

INTRODUCTION

Digital divides are among the most pressing concerns in telecommunications policy. Researchers have used a variety of methods to study the determinants and consequences of digital divides. In an overview of current research on broadband digital divides, Dwivedi and Papazafeiropoulou (2008) discuss studies using methodology from sociology, economics, strategic management, and organizational development, including case study analysis, focus group and survey methods, system dynamics and causal loop analysis, cluster analysis, and econometrics. These authors identify dozens of factors that may hinder a consumer
from adopting advanced technology, from obvious candidates such as prices and income to more subtle (and less quantifiable) influences such as attitudes toward technology.

As informative as qualitative studies may be, if attention paid to the divides is to generate light—and not just heat—then policy-makers require accurate measurement of the gaps in question. In this chapter, we assess some of the statistical tools that empirical researchers use to measure digital divides. Our focus is on econometric regression studies using data from many individuals, households, or geographic areas. Many empirical studies of the digital divide analyze a cross-section of data on the extent of digital deployment or use. Studies of broadband Internet access are a leading example (refer to the next section for citations). In these studies, researchers regress broadband subscription on characteristics of the household or the area, depending on the nature of the available data. Methods used for the binary access decision range from OLS regression, probit, and logit to more complicated estimators tailored to unique features of the data at hand (Flamm & Chaudhuri, 2007; Prieger & Hu, 2008). Researchers and policy-makers often use the results to identify subpopulations that are prone to end up on the wrong side of the digital divide.

What is missing from most of these approaches is the ability to say much about the timing of technological adoption. In this chapter we investigate the catch up hypothesis, which posits that socio-demographic groups lagging in broadband adoption eventually reach the same subscription rates as other groups. For example, take one of the results from Prieger & Hu (2008): blacks in the U.S. subscribe to broadband Digital Subscriber Line (DSL) service at a lower rate than do whites. Unanswered are the questions of whether this divide is only temporary, as predicted by the catch up hypothesis, and how rapidly the gap will close if so. These questions are close to the heart of public policy toward digital divides. If gaps exist but close quickly without intervention, policymakers may better direct public resources elsewhere. Persistent gaps, on the other hand, may warrant further study and action.

We aim to improve our understanding of appropriate techniques used to analyze the digital divide and policies aimed at reducing it. We use data on DSL adoption in the U.S. to compare the policy implications deriving from traditional cross-sectional analysis with that from duration analysis, an appropriate but under-used statistical technique in digital divide research. Our work contributes to the policy literature on the digital divide in three ways. We begin by clarifying the potential limitations of cross-sectional analysis. We also propose and explore the performance of duration analysis applied to broadband take-up data. Often data are available (or could be gathered) on how long a household has subscribed to broadband, even in cross-sectional datasets. Appropriately conducted duration analysis can then clarify the temporal dimension of the digital divide. Finally, we compare duration analysis to other methods used to examine the temporal dimension of the gap. Previous studies such as Whitacre (2007) and Flamm & Chaudhuri (2007) have analyzed data collected from different time periods. We explore whether duration analysis yields different conclusions than does panel data analysis and whether results are more easily interpretable for policy makers.

In our empirical section, we examine the demand for DSL broadband in five U.S. states. To compare traditional cross-sectional analysis with duration analysis and panel data methods, we focus attention on groups prone to the digital divide: racial minorities and women. We assess the gaps three ways. Ordinary least squares and probit regressions using the cross-sectional data, which come closest to what is done in most studies, establish a baseline for our results. We next use duration (also known as survival) analysis to speak to the pace at which gaps can be expected to close. Finally, we use the cross-sectional data, coupled with information on when households