Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Matching Platforms and HIV Incidence: An Empirical Investigation of Race, Gender, and Socioeconomic Status

Brad N. Greenwood, Ritu Agarwal

MANAGEMENT SCIENCE
Vol. 62, No. 8, August 2016, pp. 2281–2303
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2015.2232

Matching Platforms and HIV Incidence: An Empirical Investigation of Race, Gender, and Socioeconomic Status

Brad N. Greenwood
Fox School of Business, Temple University, Philadelphia, Pennsylvania 19122, greenwood@temple.edu

Ritu Agarwal
Robert H. Smith School of Business, University of Maryland, College Park, Maryland 20742, ragarwal@rhsmith.umd.edu

Although recent work has examined the adverse implications of Internet-enabled matching platforms, limited attention has been paid to whom the negative externalities accrue. We examine how the entry of platforms for the solicitation of casual sex influences the incidence rate of human immunodeficiency virus (HIV) infection by race, gender, and socioeconomic status. Using a census of 12 million patients subjected to a natural experiment in Florida, we find a significant increase in HIV incidence after platform implementation, with the largest effect accruing to historically at-risk populations (i.e., African Americans) despite documented lower rates of Internet utilization. Strikingly, our analysis reveals that HIV incidence increases in historically low-risk populations as well (e.g., individuals of higher socioeconomic status) and that men and women experience similar penalties. Identifying granular effects across subpopulations allows us to offer additional insight into the mechanisms by which matching platforms increase HIV incidence. We estimate the cumulative effect of platform entry over the five-year period of the study as 1,149 additional Floridians contracting HIV at a cost of $710 million.

Keywords: public health; two-sided matching; platforms; natural experiment; HIV; digital divide
History: Received May 13, 2014; accepted April 9, 2015, by Lorin Hitt, information systems. Published online in *Articles in Advance* November 24, 2015.

1. Introduction

The promise of public health benefits arising from the connectivity enabled by the Internet has recently captured the attention of scholars and policy makers. Frequently discussed advantages of digital connectivity for health range from support groups for chronic health conditions (Goh et al. 2016), to the use of electronic health information exchange (Agarwal et al. 2010), to applications such as Google Flu Trends that assist public health officials in monitoring epidemics (Dugas et al. 2012). However, whereas substantial research documents the benefits of connectivity, a growing literature in both information systems (Chan and Ghose 2014) and public health (Benotsch et al. 2002, Elford et al. 2001, Parsons et al. 2006) has begun to examine the pitfalls as well. One concern is that platforms that increase market efficiencies, decrease frictions, and enable the two-sided matching process for the exchange of goods and services (e.g., eBay and Amazon.com) (Bakos and Bailey 1997, Brynjolfsson and Smith 2000) may also facilitate risky behaviors, such as casual sex (Chan and Ghose 2014) and drug use (Grov 2010). Beyond the social cost of increased morbidity and mortality such behaviors create, there is also an economic burden often borne by governments and taxpayers. In light of the high costs of treatment for sexually transmitted diseases (STDs) and rehabilitation for drugs users (Schackman et al. 2006), the externalities of these sites pose significant challenges for policy makers.

Although Internet-enabled matching platforms are recognized and documented as increasing risky behavior (Benotsch et al. 2002, Chan and Ghose 2014, Kim et al. 2001, Parsons et al. 2006), and initial aggregate estimates of the effects are available, one critical issue remains understudied: a granular identification of the subpopulations to whom the negative externalities accrue. To the extent that there may be variation in the effects experienced across sociodemographic groups, and to the degree that health disparities are a significant policy concern (Secretary's Advisory Committee on National Health Promotion and Disease Prevention Objectives for 2020 2008), understanding who is vulnerable to platform entry and isolating the magnitude of this effect are important.

Much of the recent work in public health related to the use of online platforms for the solicitation of casual partners (Garofalo et al. 2007, Grov 2010) is

Management Science 62(8), pp. 2281–2303, ©2016 INFORMS

based on small-scale surveys. As a result, researchers have been unable to isolate whether the existence of the platform is increasing the total incidence of the human immunodeficiency virus (HIV) or simply displacing the spread that traditionally occurred through an off-line matching process. A notable exception to small-sample survey research is Chan and Ghose (2014), in which the authors examine the state-level spread of HIV after the introduction of Craigslist—finding a nearly 16% rise in case rate at an annual cost of approximately $65 million. However, because their analysis is conducted at the state-year level with archival incidence reports from the Centers for Disease Control (CDC), the authors are unable to isolate the socioeconomic and demographic characteristics of individuals who are penalized.

In this study we estimate the impact of Internet-enabled matching platforms on the incidence rate of asymptomatic HIV infection. We further examine which subpopulations, based on race, gender, and socioeconomic status (SES), experience greater negative externalities as a result of platform availability—a question for which an answer is not clear, ex ante. On the one hand, epidemiological accounts of HIV suggest that infection is far more prevalent among ethnic minorities and those of lower SES (CDC 2011). However, studies relating to digital inequality and variation in digital capabilities across sociodemographic characteristics underscore that these groups are modestly less likely to utilize online resources, especially so for welfare-enhancing activities (DiMaggio et al. 2004, Hargittai 2010).

We quantify the effect of platform entry on the HIV incidence rate[1] by exploiting a natural experiment: the introduction of Craigslist in the state of Florida between January 2002 and December 2006. We match the introduction of Craigslist to a census of patients in Florida hospitals. These data contain detailed information about HIV, as well as patient characteristics. This study responds to calls to examine the effects of variation in Internet use and digital capabilities across populations characterized by specific demographic and socioeconomic factors (DiMaggio et al. 2004) and, more generally, contributes to research seeking to unpack the social and economic effects of the Internet. By conducting a granular investigation of subpopulations that are disproportionately affected, we are able to generate insights into the mechanism by which matching platforms influence the rise in HIV infection (through increased hetero- or homosexual dating, a systematic change in dating patterns across groups, etc.).

From an econometric standpoint, as the entry of Craigslist is staggered temporally and geospatially, the search for sexual partners is highly localized (Zenilman et al. 1999), and our data are able to differentiate between asymptomatic and symptomatic cases of HIV, we are able to quantify the effect of platform introduction with greater precision than prior work.[2] Our identification strategy exploits the exogeneity of platform introduction into different areas at different times. Because the broader objective of Craigslist is not the facilitation of the solicitation of sexual partners (this section of the platform is non-revenue-generating and is one of the many functionalities that Craigslist offers), we can reasonably assume that the entry of the Craigslist forum is driven by its primary revenue stream (the posting of classified ads for employment) and not by the sexual proclivities of individuals in the area.

Empirical analysis yields five findings robust to multiple specifications and falsification tests. First, consistent with Chan and Ghose (2014), the introduction of Craigslist significantly increases the incidence rate of HIV infection for patients admitted to local area hospitals. Our aggregate estimate of the effect of Craigslist on asymptomatic HIV, the most proximal outcome of infection, provides further confirmatory evidence of the increased prevalence of the disease after platform entry. Estimates of the aggregate effect indicate a rise in the asymptomatic HIV infection rate of 13.5% (roughly consistent with Chan and Ghose's 15.9% estimate),[3] resulting in an annual increased treatment cost of $6.896 million in Florida as of 2006. In the aggregate, this translates to a cumulative cost of treatment of $710 million over the lifetime of all patients contracting HIV as a result of Craigslist entry during the time of the study. Second, the relative and absolute increase in HIV incidence after platform entry is highest in one subpopulation that is traditionally considered at risk for the disease: African Americans. Third, the greatest absolute increase across the income distribution is in a population not traditionally considered at risk for the disease: those of relatively higher socioeconomic status. Fourth, we find that although men are considered at greater risk for HIV than women (CDC 2011), and women are modestly less likely to use the Internet (Hargittai 2010), there is no significant difference in the effects of the platform across gender. Finally, despite the documented concentration of HIV within the Latino community,

[1] Incidence rate is defined medically as the number of new cases per population in a given time period (Dorland 2011).

[2] The archived CDC reports that have been used in prior investigations do not distinguish between levels of disease progression. Symptomatic HIV can take up to 10 years to manifest after initial infection.

[3] Without knowledge of the number of Craigslist treated hospitals by state, a further comparison is infeasible.

results suggest that the effect is not significantly different from the effect on Caucasians.

2. Background

Two-sided matching platforms that provide the digital infrastructure for distinct communities of users to transact with each other have been investigated in both the economics and information systems literatures (Bakos and Bailey 1997, Brynjolfsson and Smith 2000). Research suggests that these platforms reduce market frictions by serving as digital intermediaries between buyers and sellers (Bakos and Bailey 1997, Brynjolfsson and Smith 2000). From a transaction cost perspective, digital platforms reduce bargaining asymmetry and protect transacting agents from opportunism by increasing price transparency (Williamson 1981). From a markets and hierarchies perspective, the reduced search cost facilitated by the platforms decreases operating cost and can accelerate the buyer–seller match (Malone et al. 1987).

Whereas early manifestations of digital platforms focused on matching buyers and sellers for economic transactions, in recent years social transactions have become increasingly important. To illustrate, online dating services such as eHarmony, Match.com, and Zoosk have experienced wide success as a result of their ability to resolve frictions in off-line markets for social transactions (Bapna et al. 2012). Not unlike digital commerce platforms, dating websites facilitate the partner matching process through two mechanisms: self-selection and decreased search costs (Brynjolfsson and Smith 2000). Since many of these platforms position themselves as catering to the needs of specific demographics and market niches, ranging from eHarmony's general goal of facilitating long-term relationships to ChristianMingle.com's segmented market approach based on theological belief, they are able to efficiently subdivide the market into prospective partners who meet each other's matching criteria. Further, the platforms typically gather extensive data about users to facilitate matching, resulting in significantly reduced search costs. Simply put, by concentrating users with particular tastes or objectives on an easily searchable platform, these services reduce the search cost for acceptable matches and increase the ease of sampling (i.e., dating).

Together with sites that purportedly facilitate more enduring relationships, the Internet has also spawned platforms that enable the solicitation of casual sexual partners with no long-term relationship goal attached to them (e.g., Adult Friend Finder and Ashley Madison). Such platforms provide equivalent benefits in terms of market segmentation and reduced search cost, but they offer the added advantage of anonymity. Because users can utilize their services without fear of social stigma, anonymity decreases another significant cost: the social penalty of engaging in risky activity (Parsons et al. 2006). Although not critical for the market to function, anonymity reduces the effect of social frictions that may constrain the off-line two-sided sexual matching process (Kim et al. 2001). Such frictions, often a result of prevailing societal norms, include social castigation for promiscuity (Brown 1988), closeted sexual orientation (Floyd and Stein 2002), and even enforcement of a social contract (i.e., marriage) (MacDonald 1995). Research in psychology suggests that the security of anonymity (i.e., a lower likelihood of being discovered engaging in deviant behavior) is associated with more risky behaviors (Rogers 2010). With a guarantee of anonymity, inasmuch as individuals no longer fear social reprimand or even criminal prosecution as a result of their actions, there is likely a decreased propensity to adhere to social norms and an increased tendency to engage in risky behavior (Padgett 2007).

Unsurprisingly, empirical work in public health (Benotsch et al. 2002, Grov 2010) has highlighted the risky behavior that these sites facilitate. Not only are users more likely to engage in unprotected sex (regardless of race, as in Bingham et al. 2003; gender, as in Padgett 2007; or sexual orientation, as in Garofalo et al. 2007) with an increased number of partners (Benotsch et al. 2002, Elford et al. 2001), users of these websites are also more likely to carry an STD (Elford et al. 2001) and engage in drug use (Grov 2010). However, despite the growing research in this area, studies have yet to address the question of *who* might be more vulnerable to the adverse consequence of contracting HIV as a result of the availability of platforms that enable the solicitation of casual sex. To explore this question, we compare and contrast the characteristics of two populations: those identified as being "at risk" for HIV and those with relatively lower digital capabilities in regard to use of online resources.

2.1. Digital Disparities and Inequality

With more than 93% of the American population having access to faster-than-dial-up Internet (National Telecommunications and Information Administration/Economics and Statistics Administration (NTIA/ESA) 2013), and a large and growing segment of the population having grown up in the digital age, two increasingly common, although not entirely accurate, perceptions have emerged. First, because of the widespread availability of high-speed Internet in the United States, it has been suggested that the domestic digital divide and digital inequalities are relics of the past that further attenuate in prevalence with each passing year (Barry 2013). Second, it is commonly believed that by virtue of early socialization with technology, young people are adept technology users regardless of race, gender, and SES (Hargittai 2010). Ironically, despite the pervasiveness of these perceptions, limited empirical evidence exists to support them (NTIA/ESA 2013). Research suggests not only that racial disparities persist in Internet utilization (Hargittai 2010, NTIA/ESA 2013, Zickuhr and Smith 2012) but also that differences are largely driven by income and education inequalities (Hsieh et al. 2011). Further, even when access to online resources is equal, there exist striking differences in *how* online resources are utilized (DiMaggio et al. 2004, Hargittai 2010); i.e., there is heterogeneity in digital capabilities across subpopulations. Not only are the socioeconomically advantaged far more skillful in their exploitation of online resources, but such resources are rarely accessed and used by the socioeconomically disadvantaged for capital or welfare-enhancing activities (Hargittai 2010, Zillien and Hargittai 2009).

In light of this work highlighting the persistence of digital disparities, it is not clear to whom the negative externalities of platforms for the solicitation of casual sex will accrue. As Internet utilization rates vary across subpopulations (e.g., by race, gender, socioeconomic status) (DiMaggio et al. 2004, NTIA/ESA 2013, Zickuhr and Smith 2012), the likelihood of exploiting online resources is lower for individuals with differential familiarity and computer literacy (Hargittai 2010, Warschauer 2004, Zillien and Hargittai 2009). Because this utilization disparity usually falls along socioeconomic lines, with those of lower SES[4] and ethnic minorities suffering a disproportionate penalty (NTIA/ESA 2013, Zickuhr and Smith 2012), it is also plausible that these groups will be less likely to use online platforms for soliciting sexual partners. However, national estimates of HIV infection indicate that people living below the poverty line as well as ethnic minorities (specifically, Latinos and African Americans) are far more likely to be carriers of the HIV virus (CDC 2011). Further, epidemiological research reveals that HIV infection has an unequal effect on men (specifically bisexual or homosexual men) and individuals who engage in risky sexual behavior (CDC 2011). Thus, in comparing the two groups, those with relatively lower Internet use and digital capabilities and those at higher risk for HIV, we see a significant overlap.

The presence of this overlap suggests the need to further examine where we might expect to observe the increased HIV infection that results from the availability of matching platforms. To examine this question, we first develop numerical estimates of the number of different subpopulation members who are likely to be HIV carriers and online on the platform. Table 1 presents baseline estimates of the daily number of Craigslist users, by subpopulation, who are likely to be carriers of HIV in the state of Florida. We use available data for rates of Internet utilization and a subpopulation's likelihood of utilizing Craigslist (which has been researched extensively by groups such as the Pew Charitable Trust) and HIV incidence rates by subpopulation monitored by the CDC. Once the number of members of each subpopulation who are likely to be using the platform is determined, we create a weighted estimate of the marginal propensity to encounter an HIV-infected partner. To compute these estimates, we draw from research on the propensity of subgroups, by race, to date within and across group (Hitsch et al. 2010), and we multiply the number of HIV carriers, within subgroup, by their propensity to date members of every other subgroup. For example, the *Caucasian* dating pool, following Hitsch et al. (2010), comprises 75.53% Caucasians, 9.69% African Americans, and 14.77% Latinos.[5] From this we would conclude that 34 HIV carriers, on average, are likely to be encountered by Caucasian users of Craigslist. Because of the lack of reliable data for the relative homosexual versus heterosexual proportions represented in Craigslist users, we are unable determine the weighted propensity to encounter HIV carriers for men and women.

Two useful insights can be gleaned from these estimates. First, despite the fact that African Americans are only 15% of the population, there are likely to be more African American carriers using the platform on any given day, and more infected African Americans are likely to be encountered. Second, as expected, although the number of low SES daily users is small, their marginal probability of being HIV carriers is significantly higher than that of higher SES users.

Although the baseline estimates offer initial, approximate estimates of the number of subpopulation users who are both carriers of HIV and likely to visit Craigslist, we note that the analysis rests on two assumptions. First, the model assumes that there is an equal propensity for HIV-positive and HIV-negative individuals to solicit partners through the platform. Second, it assumes homogeneity in the relative success rate (i.e., successful solicitation of a partner) across subgroups. In light of these limitations, generalizing from the approximations may be inappropriate. Further, as noted, there is inadequate published data that can be used to compute the effect of the platform across gender. Because a detailed understanding of the impact of platform entry for

[4] According to American Psychological Association (2013), "Socioeconomic status is commonly conceptualized as the social standing or class of an individual or group. It is often measured as a combination of education, income, and occupation."

[5] Note that the propensity to date across subgroups is weighted equally across genders. Members of other subgroups (e.g., Asians, Native Americans) are dropped from the analysis and the propensity to date is normalized to 1.

Table 1 Baseline Numerical Estimates of Daily HIV Prevalence on Casual Encounters Forum of Craigslist in Florida

	Population[a]	% population online[b]	Online pop.	% using classified ads daily[c]	% using Casual Encounters[d]	Estimated pop. using Casual Encounters[e]	% HIV carriers[f]	Estimated no. of HIV carriers on Casual Encounters[g]
Caucasian	11,047,394	70	7,733,176	10	2	15,466	0.22	35
African American	2,792,980	57	1,591,999	8	2	2,547	1.72	44
Latino	3,468,988	56	1,942,633	10	2	3,885	0.59	23
Medicaid	2,312,659	49	1,133,203	6	2	1,360	2.07	28
Non-Medicaid	15,477,025	93	14,393,633	13	2	35,984	0.35	126
Men	8,681,366	66	5,729,701	12	2	13,751	0.69	94
Women	9,108,318	67	6,102,573	7	2	8,544	0.22	19
	17,789,684					81,538		368

Weighted propensity to encounter HIV carriers, by subpopulation, when soliciting partners online[h]

Subpopulation	No. of HIV carriers	Caucasian (%)	African American (%)	Latino (%)	HIV carriers encountered
Caucasian	35	75.53	9.69	14.77	34
African American	44	0.82	97.34	1.84	43
Latino	23	3.90	5.40	90.70	24
Medicaid	28				28
Non-Medicaid	126				126

[a]Gender data were obtained from Office of Economic and Demographic Research (2007). Ethnicity and SES data were obtained from Infoplease (2015). Medicaid population is proxied from percent persons living in poverty.

[b]The Latino utilization source is Fox and Livingston (2007). The The gender utilization source is Fallows (2005). All other subpopulation data came from Fox (2005). Medicaid and non-Medicaid data were proxied using income less than $30,000 and more than $75,000, respectively.

[c]Source of data is Jones (2009). Medicaid and non-Medicaid proxied as before. Point estimates used indicates the percentage of adults, within subpopulation, who use online classified ads on a daily basis as of 2009.

[d]Source of data is Quenqua (2009), which indicates that 2% of Craigslist posted ads comes from Casual Encounters in 2009. Because of data availability, we assume this number is constant across subpopulations and that the relative posting rate reflects the relative rate of Internet traffic.

[e]Estimated population calculated by multiplying the percent online population using Classified Ads Daily and the percent using Casual Encounters.

[f]Source for Medicaid/non-Medicaid data is Miguelino-Keasling (2010), proxied using HIV rate in areas with 20% households with below the poverty level. Source for all other data is Campsmith et al. (2008). Percentage derived from rate (number of people per 100,000) living with HIV.

[g]Estimated number of HIV carriers estimated by multiplying the estimated population using Casual Encounters with the percent subpopulation with HIV infection.

[h]Weighted propensity is calculated based on the propensity to date within/across subpopulations by the focal population online; source: Hitsch et al. (2010). Percent probability is equally weighted across genders and normalized to 100% across the three sampled subpopulations. Medicaid and non-Medicaid groups are assumed to sample within group because of the within-couples correlation of both income and education (Hitsch et al. 2010; Rosenfeld et al. 2011, 2014). Gender and sexual orientation are omitted because of lack of reliable estimates of the probability of individuals who identify as heterosexual and homosexual to date online.

men and women may help to isolate the mechanisms by which Craigslist increases HIV incidence, e.g., through greater homosexual or heterosexual activity, we extend our empirical analysis to also include the effect on men and women, thereby addressing the overarching question of which subpopulations, by race, socioeconomic status, and gender, experience the negative externalities of platform entry.

3. Data and Methodology

3.1. Context: Craigslist

To quantify the effect of the introduction of online platforms for the solicitation of casual sexual partners on different subsegments of the population, we exploit a natural experiment: the introduction of Craigslist into major cities in the state of Florida between January 2002 and December 2006.[6]

Craigslist is a community-based online forum for classified advertisements that was launched in 1995 in San Francisco. As of May 2014, the site, by far the largest to offer these services, received nearly 50 million visitors per month.[7] In addition to forums for the posting of résumés, employment opportunities, housing, and music shows, Craigslist also offers a personal ads section. The personal ads section, which hosts diverse opportunities for couples and groups to meet and interact, contains a forum called "Casual Encounters," which can be used for the solicitation of sexual partners. Individuals post advertisements on this forum indicating their gender and age, preferred gender of respondents, a description of other

[6] Others have similarly exploited such natural experiments (Chan and Ghose 2014, Seamans and Zhu 2014).

[7] As of January 2015, Craigslist has an Alexa ranking of 12 in the United States, thereby making it the largest platform that facilitates the solicitation of partners. Ashley Madison and Adult Friend Finder, via comparison, have Alexa ratings of 1,234 and 637, respectively.

preferences, and a method for response.[8] These advertisements, and the platform itself, satisfy each of the criteria for two-sided matching platforms: the community is populated with like-minded individuals seeking casual sexual encounters (as the title of the forum would imply), the platform is easily searchable by user preferences (i.e., age, race, and sexual orientation), and users can participate anonymously.[9]

3.2. Data

We construct a longitudinal data set that contains a census of 12 million patients admitted to hospitals in the state of Florida from January 2002 to December 2006. The source of these data is the Florida Agency for Healthcare Administration (AHCA), which provides us with bed-level information about every patient admitted into a hospital in the state during that time. As the objective of the analysis is to determine the effect of platform entry on the incidence rate of HIV, we aggregate these patient data at the hospital-quarter level so we can track the relative level of asymptomatic HIV patient admittance over time.[10] We constrain our analysis to these dates, as October 2002 is the date of the first local implementation of Craigslist (Miami-Dade), and detailed information regarding the implementation of Craigslist is unavailable after 2006.[11] Table 2 summarizes the areas where Craigslist was introduced during the study time frame together with a comprehensive list of the locations in Florida where Craigslist was eventually installed. The AHCA data set, used widely in prior research (Burke et al. 2003, 2007; Greenwood et al. 2013), offers the benefit of observing when the patient is admitted to the hospital, i.e., before or after Craigslist implementation, as well as detailed information about the patient and the medical conditions the patient has been diagnosed with through the *International Classification of Diseases, Ninth Revision* (ICD-9; National Center for Health Statistics 1998) codes associated with each admittance.

3.3. Variable Definitions

3.3.1. Dependent Variable. The dependent variable for the main analysis is the number, i.e., count, of asymptomatic HIV carriers (ICD-9 diagnosis code "V08") admitted to a specific hospital in a specific quarter. As noted earlier, we use asymptomatic, as opposed to symptomatic, HIV because the latter can take years or even decades to manifest after initial infection. The delay occurs because the body initially is capable of fighting the virus, resulting in the patient entering a clinical "latency period" 3–12 weeks after infection. A detailed explanation of disease progression is available in Pantaleo et al. (1993). This latency period, as well as the initial period of acute HIV syndrome where the patient will often experience nausea, weight loss, and swelling of the lymph nodes, is medically defined as asymptomatic HIV. Given the length of time symptomatic HIV and AIDS take to be observed, the asymptomatic stage of the disease is an appropriate and proximate effect of the Craigslist treatment. Further, since 1997, the standard treatment for HIV management has been highly active antiretroviral therapy (HAART), which can be administered in pill format and does not require hospitalization (Finzi et al. 1997). Econometrically, this mitigates significant identification problems by limiting rehospitalization (because chronic management of the disease can be conducted in an outpatient setting).

Before describing the independent variables, we make two qualifying observations. First, medically, once a patient has been classified as a symptomatic carrier, he can never be reclassified as an asymptomatic carrier.[12] Statistically, this ensures that the population of patients seeking treatment does not intermittently swap between HIV classes. Second, our dependent variable only captures patients who seek treatment for the HIV virus. Although the patient

Table 2 Florida Cities (Counties) and Craigslist Implementation Date

City	County	Implementation date
Fort Lauderdale	Broward	June 2006
Daytona Beach	Volusia	June 2006
Florida Keys	Monroe	—
Fort Myers	Lee	June 2005
Gainesville	Alachua	January 2006
Jacksonville	Duval	January 2005
Lakeland	Polk	—
Miami	Miami-Dade	October 2002
Ocala	Marion	—
Okaloosa/Walton	Okaloosa/Walton	—
Orlando	Orange	February 2004
Palm Beach	Palm Beach	April 2005
Panama City	Bay	—
Pensacola	Escambia	September 2005
Sarasota	Sarasota	June 2006
Space Coast	Brevard	—
St. Augustine	St. Johns	—
Tallahassee	Leon	June 2005
Tampa	Hillsborough	November 2003
Treasure Coast	Indian River	—

[8] See, for example, http://miami.craigslist.org/i/personals?category=cas.

[9] As of January 2015, the website requires posters to register a telephone number to submit an ad. These user protections had not been implemented during the time of our study.

[10] The quarter level is the most granular time window available from the AHCA.

[11] Per http://www.craigslist.org/about/expansion (accessed May 13, 2014).

[12] The stages of HIV are defined by the CDC based on CD4 T-cell count (CDC 2008), which will only decrease after the initial period of acute HIV syndrome.

may be diagnosed at any number of locations (e.g., blood drives, annual visits to a primary care provider, mobile testing units), protocol suggests that an individual (upon receiving an initial diagnosis at a blood drive or clinic) will be referred to a hospital for more thorough antibody/antigen testing. This is done both to confirm that the patient is not a carrier of a drug-resistant form of HIV and to ensure that guidelines for the use of antiretroviral drugs are followed (Panel on Antiretroviral Guidelines for Adults and Adolescents 2006, Little et al. 2002).

3.3.2. Independent Variables. The central independent variable of interest is the dichotomous indicator *Craigslist*, which indicates that the focal patient has been admitted to a hospital that is in the same county as a city in which Craigslist has been released (as determined by Table 2).[13] Because our data from AHCA areaggregated at the quarter level, this variable is set to 1 in the first *full* quarter during which the county has had access to Craigslist, as well as all subsequent quarters.

To investigate the effect of Craigslist introduction across various different subpopulations, we disaggregate our dependent variable, *Total Count*, based on the various sociodemographic indicators that our data capture. The first, race/ethnicity (henceforth referred to as *race*), comprises three groups: African American, Latino, and Caucasian. We next subdivide our data based on the gender of the patient, as casual sex-seeking behavior is observed more often in men (Elford et al. 2001). Finally, as a proxy for SES, we classify patients based on whether or not they are recipients of Medicaid. Because Medicaid is available only to people with an income of $1,869 a month or less ($1,439 in 2002 dollars) in the state of Florida,[14] it is a reasonable indicator of low SES. Other patients are aggregated to the "non-Medicaid" group and therefore represent individuals of middle and high SES.

To execute the difference in difference strategy, we further include fixed effects for the hospitals[15] as well as fixed effects for the 20 time quarters of the study period. Hospitals that are absent from a significant portion of the study (37 of 260), because they either open after halfway through the sample, close before the halfway point of the sample,[16] or have no record of admitting an HIV-positive patient (in which case there is no variation in the dependent variable), are dropped from the sample. With these restrictions, the final data set consists of 223 hospitals over 20 quarter time periods. Summary statistics and correlations are reported in Table 3.

3.4. Empirical Strategy

As our dependent variable is a count, we use a fixed effect negative binomial estimation with robust standard errors clustered on the county, resulting in an estimation of the aggregate difference in the differences between treated and untreated hospitals. We cluster our standard errors at the county level because the treatment, *Craigslist*, is applied at the county level.[17] We model the number of asymptomatic HIV patients admitted into the focal hospital j at time t, (y_{jt}), using the following specification:

$$y_{jt} = \beta_1 s_1 + M'\theta_1 + X'\delta_1 + \nu + \varepsilon.$$

The variable s_1 is the dichotomous indicator of Craigslist treatment; M is the vector of hospital fixed effects, and X is the vector of time fixed effects. The term $\{\beta_1\}$ is the parameter to be estimated, and ν represents the constant. As discussed previously, after estimating the cumulative effect for the entire population, i.e., the aggregate difference in difference, we estimate the effect on the population based on race (Caucasian, African American, or Latino), SES (Medicaid or non-Medicaid), and gender (male or female). Results are available in Table 4.

4. Results

4.1. Base Model

In Table 4 we first note a strong and significant increase in the number of asymptomatic HIV patients admitted after the entry of Craigslist, consistent with the effects reported in Chan and Ghose (2014). From an economic perspective, this translates to an absolute increase of 1.36268 patient admittances per hospital quarter with asymptomatic HIV (or a relative increase of 13.45% over the average 10.13116 patients per hospital quarter admitted in the absence of Craigslist).[18]

[13] The Craigslist "treatment" is applied to the county within which the city is located.

[14] See http://www.floridamedicaideligibility.com/asset.html (accessed May 13, 2014).

[15] Although the treatment of our study is applied at the county level, we include hospital fixed effects to account for the fact that patients of different backgrounds may select into seeking treatment at different hospitals within the same county. One alternative empirical strategy would be to execute the difference-in-differences model at the county-quarter level. However, because there are significantly more hospitals than counties, executing the difference-in-differences estimation at the hospital-quarter level increases the ability to control for unobserved heterogeneity in the empirical estimations.

[16] Robustness checks have been conducted eliminating all hospitals not present for the study's duration. Results are consistent and available upon request.

[17] We validate the assumption of spatial reach of treatment at the county level using alternative radii in robustness checks described in §5.1.

[18] All marginal effects are calculated using the default `margins` command of Stata 12.1. This is done by calculating the average predicted change in the dependent variable after the treatment, i.e., Craigslist entry, is manipulated. All other covariates are fixed at the mean. Standard errors for the post-estimation command are obtained through the delta method.

Table 3 **Summary Statistics ($N = 4{,}349$)**

Variable	Mean	Std. dev.	Min	Max	1	2	3	4	5	6	7	8
1 *Craigslist*	0.194	0.395	0	1								
2 *Total Count*	10.582	22.762	0	246	0.2083							
3 *Total Caucasian*	3.313	5.636	0	63	0.1237	0.8170						
4 *Total African American*	5.757	15.318	0	171	0.1774	0.9704	0.6950					
5 *Total Latino*	1.260	4.073	0	55	0.3051	0.7370	0.5030	0.6367				
6 *Total Medicaid*	3.631	9.759	0	138	0.1849	0.9567	0.6878	0.9609	0.7170			
7 *Total Non-Medicaid*	6.950	13.722	0	141	0.2140	0.9783	0.8660	0.9263	0.7126	0.8757		
8 *Total Male*	6.079	12.503	0	120	0.2226	0.9710	0.8603	0.9063	0.7572	0.8865	0.9802	
9 *Total Female*	4.503	11.033	0	128	0.1775	0.9626	0.7105	0.9750	0.6624	0.9691	0.9075	0.8700

Table 4 **Negative Binomial Estimates of Total Number of Asymptomatic HIV Cases by Subpopulation Asymptomatic HIV Incidence, 2002–2006**

	Dependent variable							
	(1) *Total Count*	(2) *Total Caucasian*	(3) *Total African American*	(4) *Total Latino*	(5) *Total Non-Medicaid*	(6) *Total Medicaid*	(7) *Total Male*	(8) *Total Female*
Craigslist	0.126***	0.0781	0.146***	0.158	0.120**	0.130***	0.120*	0.134***
	(0.0349)	(0.0509)	(0.0327)	(0.0968)	(0.0427)	(0.0300)	(0.0477)	(0.0320)
Constant	4.012***	2.263***	3.793***	−0.241*	3.401***	3.229***	3.163***	3.449***
	(0.0278)	(0.0426)	(0.0476)	(0.116)	(0.0400)	(0.0434)	(0.0438)	(0.0303)
Hospital fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log pseudolikelihood	−9,085.13	−6,598.22	−6,624.08	−3,480.73	−8,080.41	−5,819.21	−7,621.38	−6,562.2

Notes. Number of observations = 4,349. Robust standard errors are in parentheses (clustered on county).
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Further, there is a strong and significant effect across each of the subpopulations (with the exception of Latinos, with a p-value of 0.103, and Caucasians, with a p-value of 0.125). We note that these marginally significant effects are positive (as a negative effect on the asymptomatic HIV rate would be counterintuitive).

In row (1) of Table 5, we tabulate the marginal effects from Table 4. Consistent with the baseline numerical estimates, the largest effect is within the African American population (0.856729 patients per hospital quarter, or a 15.7% increase over those who are admitted in the absence of Craigslist).[19] Further, as suggested by the numerical estimates, the largest absolute effect is among the non-Medicaid population, with an increase of 0.850716 (12.7%) versus 0.48142 (13.9%) patients per hospital quarter. Finally, we see a modestly larger effect for men than women: 0.741839 (12.7%) versus 0.616944 (14.3%), respectively.[20] Note that all percentage increases are the percent change over those patients, within subpopulation, who are admitted in the absence of Craigslist. All else equal, these findings suggest that for race, 62.9% of the total effect accrues within the African American community; for SES, 62.4% of new cases are from the non-Medicaid population; and across gender, 54.4% of new cases are men. Results from an ordinary least squares (OLS) estimation (see Table 6) corroborate these findings (with African Americans and those of higher SES experiencing the largest within-subpopulation effect). Further, in this estimation we see a significant effect for Caucasians and Latinos.

4.2. Log Model of Relative Effect

Whereas our initial estimates provide some insights into which subpopulations are being affected to the greatest *absolute* degree, i.e., the total increase in the number of patients, the estimates of the relative change of the incidence rate, i.e., percent change, in each of these subpopulations presented above should be interpreted with caution because we cannot statistically establish if the percent change across models is significantly different. To compare the *relative* effect sizes across models, we next take the log (+1) of each

[19] Results from a seemingly unrelated regression (SUR) indicate that the effect sizes for men and women are not significantly different. Further, the SUR indicates that the effect on African Americans is significantly larger than either Caucasians or Latinos, and the effect on high SES is significantly larger than that on low SES.

[20] To ensure the robustness of the results, we further execute this analysis using a zero-inflated negative binomial regression (inflating both by the number of patients who have been diagnosed with asymptomatic HIV at the hospital to date and by the number of patients who have been diagnosed with all levels of the HIV virus). Results of the Vuong test indicate no relative advantage to either approach (negative binomial or zero-inflated negative binomial). All other results remain consistent and are available upon request.

Table 5 Tabulation of Marginal Effects Using Various Estimators

	Subpopulation							
Estimator	All	Caucasian	African American	Latino	Non-Medicaid	Medicaid	Male	Female
(1) Negative binomial	1.3627*	0.2641	0.85763*	0.19729	0.85072*	0.48142*	0.74184*	0.61694*
(2) Base OLS	2.8940*	0.6010*	1.64800*	0.58900*	1.91500*	0.97900*	1.59900*	1.29500*
(3) Log model	12.70%*	5.810%*	13.900%*	9.070%*	12.300%*	10.000%*	13.200%*	11.400%*
(4) Rel time$_{(t+4)}$	1.8663*	0.5199*	0.92814*	0.42904*	0.86764*	0.96771*	0.88813*	0.92244*
Change in the patient's marginal probability of diagnosis (reported in basis points)								
(5) Logit[a]	3.782*	1.474	16.112*	2.981	3.104*	7.615*	4.184*	3.574*
(6) LPM[b]	5.334*	2.105*	14.664*	3.118	4.605*	7.972*	5.942*	4.857*

[a]Marginal effects for the logit model identifies the mean change based on the mean of other variables. The net size of this effect will change based on the location across the probability distribution (as determined by other independent variables).
[b]The marginal effects from the LPM are reported based on the fully interacted model between the Craigslist treatment, race, gender, and SES.
*Indicates marginal effect is significant at the $p < 0.05$ level.

Table 6 OLS Estimates of Total Number of Asymptomatic HIV Cases by Subpopulation Asymptomatic HIV Incidence, 2002–2006

	Dependent variable							
	(1) *Total Count*	(2) *Total Caucasian*	(3) *Total African American*	(4) *Total Latino*	(5) *Total Non-Medicaid*	(6) *Total Medicaid*	(7) *Total Male*	(8) *Total Female*
Craigslist	2.894***	0.601**	1.648***	0.589***	1.915***	0.979***	1.599***	1.295***
	(0.335)	(0.200)	(0.279)	(0.169)	(0.362)	(0.219)	(0.284)	(0.192)
Constant	64.93***	10.51***	52.97***	0.636***	34.83***	30.10***	27.10***	37.83***
	(0.333)	(0.149)	(0.325)	(0.101)	(0.279)	(0.214)	(0.261)	(0.194)
Hospital fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.961	0.858	0.955	0.899	0.937	0.941	0.938	0.946

Notes. Number of observations = 4,349. Robust standard errors are in parentheses (clustered on county).
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

of the dependent variables in our base model and replicate our analysis. The estimator is an OLS.

Results of the log model shown in Table 7 provide further evidence of the effect of Craigslist entry. First, confirming the OLS estimations, we see a consistent effect both in the aggregate and in the individual subpopulations (including Latinos and Caucasians). Furthermore, we see that not only is the absolute effect on the African American community the largest, but the relative effect is the largest as well. The SUR confirms that the effect sizes are statistically different. Interestingly, however, results of the SUR continue to indicate that the effect size between Latinos and Caucasians is insignificantly different. Strikingly, we also see the largest *relative* effect across socioeconomic groups is within the non-Medicaid population (although the SUR indicates that the effect sizes are not significantly different). Finally, the SUR indicates that the effect sizes across gender are not significantly different. See row (3) of Table 5 for a tabulation of the marginal effects by subpopulation.

4.3. Relative Time Model

Our core empirical strategy has been an aggregate difference in the differences approach between treated and untreated hospitals. For this estimate to be valid, one concern that must be addressed is whether or not there is homogeneity in the pretreatment trend between treated and nontreated hospitals. This concern arises because it is possible that unobservable and randomly distributed environmental factors, which are native to the individual metropolitan areas, are causing heterogeneity in the pretreatment asymptomatic HIV incidence rate. For example, because Craigslist enters larger metropolitan areas first (e.g., Miami) it is possible that there is a different trend in the HIV rate between these areas and those that do not receive the treatment. To rule out this potential confound and further substantiate the claim that the entry of Craigslist can be treated as an exogenous event, we execute the following estimation:

$$y_{jt} = \rho'[s_2 \cdot \varphi] + M'\theta_2 + X'\delta_2 + \nu + \varepsilon,$$

where y_{jt} is the number of asymptomatic carriers admitted to hospital j during time t, s_2 is a dichotomous variable that indicates whether or not Craigslist treatment will ever affect hospital j during the study, and φ is a series of time dummies that indicate the relative chronological distance between time t and Craigslist implementation at hospital j, i.e., the

Table 7 OLS Estimates of Log Number of Asymptomatic HIV Cases by Subpopulation Asymptomatic HIV Incidence, 2002–2006

	Dependent variable							
	(1) log(*Total Count*)	(2) log(*Total Caucasian*)	(3) log(*Total African American*)	(4) log(*Total Latino*)	(5) log(*Total Non-Medicaid*)	(6) log(*Total Medicaid*)	(7) log(*Total Male*)	(8) log(*Total Female*)
Craigslist	0.127***	0.0581*	0.139***	0.0907*	0.123**	0.100***	0.132***	0.114**
	(0.0318)	(0.0251)	(0.0359)	(0.0342)	(0.0354)	(0.0288)	(0.0364)	(0.0397)
Constant	4.045***	2.341***	3.873***	0.513***	3.444***	3.337***	3.238***	3.517***
	(0.0359)	(0.0267)	(0.0369)	(0.0302)	(0.0317)	(0.0334)	(0.0302)	(0.0371)
Hospital fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.881	0.788	0.870	0.746	0.859	0.833	0.851	0.842

Notes. Number of observations = 4,349. Robust standard errors are in parentheses (clustered on county).
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

number of quarters preceding, or following, the local implementation of Craigslist. The vector M represents the vector of hospital fixed effects, and X is the vector of time fixed effects. The vector $\{\rho\}$ contains the parameters to be estimated, and ν represents the constant. Intuitively, this model allows us to determine whether or not there is a pretreatment trend that is disproportionately affecting Craigslist hospitals, as opposed to non-Craigslist hospitals. As in our previous analyses, after estimating the cumulative effect for the entire population, we estimate the effect for each subpopulation based on race (Caucasian, African American, or Latino), SES (Medicaid or non-Medicaid), and gender (male or female). The estimator is a negative binomial with robust standard errors clustered on the county, and the omitted relative time variable is $Craigslist_{t0}$.

Results, shown in Table 8, are consistent with previous findings and rule out several possible econometric concerns related to the assumptions of the difference in difference strategy. First, we see no effect of the Craigslist treatment on the asymptomatic HIV incidence rate before it is implemented (both in the aggregate and across subpopulations), with one exception. Of the 32 observed pretreatment quarters, one is significant (*Male Count* in $Craigslist_{t-3}$).[21] However, the coefficient is negative and nine months prior to implementation with no discernible trend in the remainder of the pretreatment dummies. From this we conclude that the assumption of the difference-in-differences models, i.e., that there is no significant and detectable dissimilarity in disease incidence trends pretreatment, is not being violated (Angrist and Pischke 2008). Second, we see (in column (1)) that the effect of Craigslist is insignificant initially but grows in magnitude over time, as we might expect, becoming significant roughly nine months after implementation. Third, as seen in column (3) and confirming previous estimates, the largest and fastest effect to manifest is within the African American population, becoming significant six months after implementation, with a net effect of 0.97 patients per hospital quarter (a 16.9% increase) one year after implementation. Interestingly, we also see a significant effect within the Caucasian and Latino populations, albeit taking substantially longer (six months) to become observable. All else equal, this suggests that the initial estimations using the negative binomial may be misleading because the early dormant period is masking the later significant effect on these subpopulations.[22] We also see that, despite the fact that the non-Medicaid population experiences the largest effect, the Medicaid population's effect is faster to manifest. We observe this same counterintuitive effect for women (although, once again, the marginal effects across gender are insignificantly different). See row (4) of Table 5 for a tabulation of the marginal effects by subpopulation at relative time $t+4$.

4.4. Marginal Probability of Diagnosis

One further plausible explanation of the change in HIV incidence rate is that although there is a net increase in the number of diagnoses, the *marginal* effect, per patient, is not changing. As a result, very small changes in the population of each county, which are statistically unidentifiable, may be driving the effect. To the degree that HIV affects a small percentage of the American population (1.1 million of the 295.5 million population in 2005[23]), we must ensure that the entry of Craigslist is not simply influencing the number of diagnoses, but the marginal likelihood that the patient is a carrier as well, which should be independent of the population changes

[21] We also note that the pretreatment dummy (not displayed) for *Latino Count* in $Craigslist_{t-5}$ is negative and significant. Given the number of coefficients that have been estimated, the probability of 2 out of 40 being correlated is 5%, our significance cutoff level.

[22] We thank the associate editor for this insight.

[23] See http://www.cdc.gov/mmwr/preview/mmwrhtml/su6001a19.htm (accessed May 13, 2014).

Table 8 Relative Time Negative Binomial Estimates of Number of Patients Admitted, with Asymptomatic HIV Quarter-Level Relative Time Fixed Effects Longer Than One Year from Implementation Date Omitted

	Dependent variable							
	(1) *Total Count*	(2) *Total Caucasian*	(3) *Total African American*	(4) *Total Latino*	(5) *Total Non-Medicaid*	(6) *Total Medicaid*	(7) *Total Male*	(8) *Total Female*
Craigslist$_{t-4}$	−0.0164	0.0229	−0.0591	0.0498	−0.0240	−0.0123	−0.0305	−0.00849
	(0.0415)	(0.0395)	(0.0682)	(0.0777)	(0.0488)	(0.0450)	(0.0538)	(0.0335)
Craigslist$_{t-3}$	−0.0507	−0.0338	−0.0547	−0.0627	−0.0397	−0.0606	−0.0762**	−0.0175
	(0.0316)	(0.0567)	(0.0369)	(0.127)	(0.0362)	(0.0483)	(0.0254)	(0.0588)
Craigslist$_{t-2}$	−0.0318	−0.0758	0.0134	0.0163	−0.0287	−0.00499	−0.0543	0.0250
	(0.0477)	(0.0600)	(0.0614)	(0.101)	(0.0456)	(0.0500)	(0.0477)	(0.0510)
Craigslist$_{t-1}$	0.0364	0.0502	0.0283	0.0410	0.0392	0.0359	0.0590	0.00845
	(0.0381)	(0.0460)	(0.0341)	(0.124)	(0.0465)	(0.0454)	(0.0407)	(0.0496)
Craigslist$_{t0}$	Omitted group							
Craigslist$_{t+1}$	−0.0121	−0.0610	−0.0194	0.172	0.0132	−0.0265	0.00181	0.00392
	(0.0364)	(0.0422)	(0.0536)	(0.0880)	(0.0447)	(0.0299)	(0.0498)	(0.0494)
Craigslist$_{t+2}$	0.104	0.0857	0.119*	−0.00876	0.0547	0.183**	0.0692	0.151**
	(0.0551)	(0.0707)	(0.0560)	(0.105)	(0.0587)	(0.0618)	(0.0684)	(0.0487)
Craigslist$_{t+3}$	0.114***	0.0275	0.121***	0.114	0.112**	0.106*	0.113***	0.0911*
	(0.0299)	(0.0396)	(0.0269)	(0.150)	(0.0347)	(0.0465)	(0.0335)	(0.0416)
Craigslist$_{t+4}$	0.163***	0.146*	0.150***	0.296***	0.118**	0.238***	0.137**	0.187***
	(0.0286)	(0.0569)	(0.0395)	(0.0560)	(0.0374)	(0.0305)	(0.0519)	(0.0447)
Constant	4.015***	2.261***	3.801***	−0.293**	3.405***	3.225***	3.168***	3.446***
	(0.0311)	(0.0406)	(0.0425)	(0.0967)	(0.0440)	(0.0488)	(0.0472)	(0.0350)
Hospital fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log pseudolikelihood	−9,082.06	−6,594.32	−6,622.78	−3,476.80	−8,080.30	−5,811.76	−7,618.18	−6,560.03

Notes. Number of observations = 4,349. Robust standard errors are in parentheses (clustered on county).
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

pre- and post-Craigslist entry. We therefore conduct a bed-level investigation to determine the increase in the marginal probability that a patient who is admitted to a Craigslist treated hospital is a carrier of the HIV virus. The dependent variable for this analysis is dichotomous and indicates whether or not the focal patient is an asymptomatic HIV carrier.

We first conduct the bed-level analysis using a logit model. However, because of some well-known concerns with logit models, we further use a linear probability model (LPM) to ensure robustness of the results. As noted by many researchers (Ai and Norton 2003, Hoetker 2007, Zelner 2009), the interpretation of interaction terms in logit models, i.e., the interaction between the *Craigslist* treatment variable and the patient-specific subpopulation indicators, is complex and requires the simulation of the marginal effects post estimation because the effect of changes in the independent variable of interest is dependent on the values of other covariates in the model. As the statistical tools that have been designed both by Zelner (2009) and by Ai and Norton (2003) have been developed for only a single interaction term to be analyzed (thereby ignoring concomitant changes in the other interaction terms), the use of interaction terms for a bed-level analysis will lead to a biased estimate of the marginal effect. We therefore conduct our initial bed-level analysis on the aggregate population and repeat it for each of the subpopulations individually.[24] Results are in Table 9, and tabulation of the marginal effect is shown in row (5) of Table 5. As in the main results, we find a significant effect across SES and gender[25] and for African Americans, with nonsignificant coefficients for the Latino and Caucasian populations.

As a result of challenges with interaction terms in logit models, we further estimate the marginal increase in the likelihood of HIV diagnosis after treatment with Craigslist using an LPM. Although the LPM allows for a meaningful interpretation of the interaction terms between Craigslist and each of the different indicators of patient characteristics (i.e., the subpopulations to which the focal patient belongs), it is not without flaws. First, it can introduce heteroscedasticity into the estimates, a concern

[24] Note also that the rarity of HIV infection (less than 0.5% of the patients in our sample) can lead to a biased estimation of the standard errors (King and Zeng 2001). To mitigate this concern, we have also estimated our results using a rare events logistic regression. The results remain consistent and are available upon request.

[25] Because the marginal effect of the coefficients will depend on the values of the other covariates in the model, directly comparing coefficients across the model is not appropriate.

Table 9 Logit Estimates of the Marginal Increase in Likelihood of Patient Infection with Asymptomatic HIV (Dependent variable = *HIV Carrier* (0/1))

	Population sample							
	(1) Total	(2) Total Caucasian	(3) Total African American	(4) Total Latino	(5) Total Medicaid	(6) Total Non-Medicaid	(7) Total Male	(8) Total Female
Craigslist	0.0988**	0.0816	0.121**	0.0962	0.0976***	0.1023**	0.0823*	0.123**
	(0.0315)	(0.0476)	(0.0392)	(0.0774)	(0.0241)	0.0325802	(0.0354)	(0.0403)
Constant	−4.794***	−5.818***	−4.256***	−5.499***	−4.3589***	−5.0324***	−4.811***	−4.789***
	(0.0342)	(0.0476)	(0.0429)	(0.0757)	(0.0406)	(0.0382)	(0.0385)	(0.0449)
Hospital fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log pseudolikelihood	−272,783.24	−98,755.98	−126,356.59	−35,095.26	−83,073.007	−186,044.67	−150,102.52	−119,586.56
Observations	11,793,617	8,090,330	1,867,208	1,747,153	1,967,303	9,803,373	5,075,571	6,646,988

Note. Robust standard errors are in parentheses (clustered on county).
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

that we address by using heteroscedastic-consistent Huber-White standard errors clustered by county. Second, the LPM can produce estimates that exist outside the [0..1] bounds of the model. A postestimation inspection shows that predicted values remain with the interval. *Total Caucasian* serves as the base case for race, *Total Non-Medicaid* patients serve as the base case for SES, and *Female* as the base case for gender. Results of the LPM (see Table 10) add interesting nuance to the main findings. Whereas our previous analyses had suggested that there is no significant difference in the effect across gender, these results indicate that there is a slightly larger effect for men. Further, results confirm that there is no significant difference in the *relative* penalty across SESs. Finally, the strikingly consistent finding across all estimations is also evident in the LPM: results indicate that a large penalty accrues to the African American population.[26]

Table 10 LPM Estimates of the Marginal Increase in Likelihood of Patient Infection with Asymptomatic HIV (Dependent Variable = *HIV Carrier* (0/1))

	(1)	(2)	(3)
Craigslist	0.000533***	0.000527***	−0.000542
	(0.000103)	(9.14e−05)	(0.000498)
Total African American		0.00920***	0.00836***
		(0.000959)	(0.000915)
Total Latino		−0.000638	−0.000212
		(0.000652)	(0.000420)
Total Gender		0.00240***	0.00205***
		(0.000387)	(0.000338)
Total Medicaid		0.00272***	0.00272***
		(0.000477)	(0.000560)
Total African American × Craigslist			0.00304*
			(0.00126)
Total Latino × Craigslist			−0.000571
			(0.000326)
Total Gender × Craigslist			0.00139*
			(0.000583)
Total Medicaid × Craigslist			−1.09e−07
			(0.000718)
Constant	0.00921***	0.00271**	0.00290***
	(0.000116)	(0.000799)	(0.000752)
Hospital fixed effects	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes
R-squared	0.005	0.009	0.009

Notes. Number of observations =11,793,617. Robust standard errors are in parentheses (clustered on county).
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

4.5. Results Summary

As indicated earlier, marginal effects from each of the estimations are summarized in Table 5. We highlight several findings that run counter to the baseline numerical estimates (see Table 1) with respect to where the negative externalities of Craigslist accrue. For the subpopulation characteristic of race, results of the empirical analysis corroborate the numerical estimates. Although there is an effect experienced by each of the races (the relative time model suggests that the initial dormant period may be masking the significant effect on the Latino and Caucasian subpopulations), the largest absolute and relative effect is within the African American community and is of considerable magnitude. However, there appears to be no statistically significant difference in the effect across Latinos and Caucasians in either relative or absolute numbers (which is surprising, given the estimated number of HIV-infected Latinos using Craigslist as a proportion of the Latino population). With respect to SES, findings also add interesting nuance to the numerical estimates. Results indicate that, across the two categories of SES, the largest absolute effect is within the non-Medicaid subpopulation, a group not traditionally considered at risk for the disease, although the effect accrues faster for Medicaid patients, and the relative effect size (i.e., percent increase) is not significantly different. Finally, although we were unable

[26] Note that the low model fit (R^2) for the LPM is consistent with extant epidemiological research (Heinzl et al. 2005).

to obtain reliable baseline estimates for the effects on men and women, the findings for gender challenge conventional wisdom. Despite the higher risk of HIV infection among men (CDC 2011) and the fact that women are marginally less likely to be Internet users (Hargittai 2010), there appears to be no significant difference in the accrual of the disease across gender (although there is a modestly higher effect for men in relative diagnosis).

5. Robustness Checks

Whereas the main results in the paper have been stable across different specifications, consistently pointing to significantly varying effects across sociodemographic variables, several alternative explanations of the effect exist. We therefore conduct an extensive set of falsification tests to further eliminate competing explanations. With the exception of the treatment radius, where we validate the radius of the treatment effect,[27] and changes in the propensity to be tested for HIV, where data are unavailable, we examine the robustness of our findings at both the aggregate and subpopulation levels.

5.1. Varying Treatment Radius

Our first empirical concern is validating the size of the treatment radius, i.e., the geospatial effect of Craigslist. Because of the relative mobility of Floridians, a state with 16.05 million registered automobiles[28] for a population of 18.09 million people, we coded the treatment at the county level for the main analysis. To ensure robustness of the assumption, we further investigate whether contracting or enlarging the treatment radius, i.e., the geographic area assumed to be affected when Craigslist is implemented, changes the size of the net effect. From a policy perspective, a deeper understanding of spatial reach is important. If, for example, hospitals within the city limits of Miami and Tampa were unaffected by Craigslist implementation, but neighboring suburban hospitals were affected, this would change the policy implications of the result substantially. To conduct this analysis, we introduce two new dummy variables: *Craigslist* (*City*) and *Craigslist* (*Greater County*). *Craigslist* (*City*) indicates that the hospital resides within the limits of a city that has received the Craigslist treatment. *Craigslist* (*Greater County*) indicates that the hospital is within a contiguously adjacent county to the county that received the Craigslist treatment. Results are available in Table 11. As expected, a smaller treatment radius (at the city level; column (1) of Table 11) results in a significant change in the asymptomatic HIV infection rate. However, consistent with research showing that search patterns for sexual partners are highly localized (Zenilman et al. 1999), the effect becomes insignificant as the radius increases beyond the county border (column (3) of Table 11). These results support the validity of the assumptions that the geographic dispersion of Craigslist implementations is sufficient to be considered a natural experiment, the implementation of Craigslist has a significant impact on the asymptomatic HIV infection rate within cities and local areas that are treated, and it is appropriate to cluster estimates at the county level.

Table 11 Negative Binomial Estimates of Introduction of Craigslist at Different Geographic Radii, Where *Craigslist* (*City*) Indicates Introduction of Platform to the Focal City and *Craigslist* (*Greater County*) Indicates Introduction of Platform to the County and All Bordering Counties (Dependent Variable = *Total Count*)

	(1)	(2)	(3)
Craigslist (City)	0.143*** (0.0295)		
Craigslist		0.126*** (0.0349)	
Craigslist (Greater County)			−0.0484 (0.0449)
Constant	3.986*** (0.0283)	4.012*** (0.0278)	3.998*** (0.0349)
Hospital fixed effects	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes
Log pseudolikelihood	−9,087.41	−9,085.13	−9,095.56

Notes. Number of observations = 4,349. Robust standard errors are in parentheses (clustered on county).
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

5.2. Other Diagnoses

One further concern with our initial estimations is that patients are not necessarily being admitted into hospitals because of a suspected infection with the HIV virus but, rather, for some other condition. For example, it is possible that a patient who suffers a heart attack will be admitted to the hospital and, during the course of her treatment, be diagnosed with HIV. Similarly, it is possible that a patient who has already been diagnosed with HIV is admitted to the hospital because of a heart attack. Because the Florida AHCA data do not allow us to track patients over time (for privacy concerns), this raises the possibility that the dependent variable may be counting patients multiple times. Although it is unlikely that patients who have been diagnosed with asymptomatic HIV will be readmitted to the hospital because of HIV-related illness (unless they have been infected with an opportunistic disease, in which case they will subsequently be diagnosed with symptomatic HIV and not be observed in

[27] Results from the treatment radius analysis for subpopulations are broadly consistent and available upon request.

[28] http://www.fhwa.dot.gov/policy/ohim/hs06/htm/mv1.cfm (accessed October 28, 2015).

Table 12 Negative Binomial Estimates of Effect of Craigslist Introduction on Patients Affected by Other Conditions

	Dependent variable					
	(1) *AMI*	(2) *Stroke*	(3) *Melanoma*	(4) *Lung cancer*	(5) *Symptomatic HIV*	(6) *Osteosarcoma*
Craigslist	0.00848	−0.0145	0.0416	0.0188	−0.0388	−0.0197
	(0.0266)	(0.0386)	(0.0712)	(0.0217)	(0.0332)	(0.0850)
Constant	4.979***	4.753***	0.893***	4.352***	4.749***	1.456***
	(0.0185)	(0.0219)	(0.118)	(0.0169)	(0.0196)	(0.114)
Hospital fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Log pseudolikelihood	−16,282.17	−14,047.86	−4,521.84	−13,438.72	−12,364.11	−3,784.098

Notes. Number of observations = 4,349. Robust standard errors are in parentheses (clustered on county).
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

the sample),[29] it is possible that a patient who has been previously diagnosed is readmitted for an emergent condition (such as a heart attack).

To mitigate this alternative explanation, we reexecute our estimations using the number of patients afflicted with other noncommunicable medical conditions, which should be uncorrelated with Craigslist introduction, as the dependent variable. The rationale for this approach is twofold. First, if patients are being admitted randomly because of other conditions, i.e., the admittance is uncorrelated with the Craigslist treatment, then the effect of additional admittances will be wholly contained within the error term and not bias the estimation of the asymptomatic HIV effect. Second, we use noncommunicable diseases because if, as we argue, Craigslist is increasing the density of social connections, then it may influence the incidence rate of communicable diseases (e.g., influenza, pneumonia). Empirically, we use the four leading causes of death in the United States (heart attack (AMI), stroke, melanoma, and lung cancer; see CDC 2013). In addition, we test the effect of the Craigslist treatment on symptomatic HIV to ensure that Craigslist is not positively influencing the incidence rate during the time of our study. Finally, because HIV is a relatively rare condition, and it is concentrated within ethnic minorities and those of low SES, we replicate the results with a form of cancer that exhibits these same properties: osteosarcoma (i.e., bone cancer) (Ottaviani and Jaffe 2010).[30] Results are available in Table 12.

In each of these estimations we see that the entry of Craigslist into a county is not significantly influencing the incidence rate of the measured conditions. Econometrically, this suggests that the effect of patient readmission is not biasing the estimate of the effect of Craigslist on the asymptomatic HIV rate. Furthermore, it eliminates the possibility that the increased admittance rate of HIV patients is a result of structural changes occurring at Craigslist hospitals. For instance, if local or state governments had been making large-scale capital investments in their local area hospitals, then patients with more severe conditions, such as HIV, may travel to these hospitals (because of their increased ability to treat severely ill patients). Findings suggest that this alternative explanation is not driving the effect. Finally, our results corroborate that there is no effect on the symptomatic HIV rate.[31]

5.3. Exposure Model

Whereas our estimates of the effect of Craigslist on noncommunicable diseases suggest that there is no aggregate change in the population after the entry of Craigslist, and our estimation of the pretreatment trend in the relative time model indicate that there is no change in HIV incidence pretreatment (either in the aggregate or in the individual subpopulations), it is plausible that there are changes in the *relative* composition of the population between treated and untreated counties, after treatment, which may be driving the effect. In other words, there are changes in the demographic makeup of the population in each county, such as widespread migration of African Americans to larger cities. These changes may be the result of any number of social factors ranging from gentrification to the economic depression and bottoming out of financial markets in 2002 (a delayed effect of the dot-com bubble burst).

To eliminate this potential confound, we estimate an exposure model where we control for the population of the county in which the hospital is located,

[29] As is well established in the virology literature (Palella et al. 2006), treatment of asymptomatic HIV in the long term typically occurs in an outpatient setting. Indeed, the definitive source for tracking HIV is the HIV Outpatient Study (Bozzette et al. 1998, Holmberg et al. 2004).

[30] We thank the anonymous reviewer for this suggestion.

[31] These estimations have also been performed on the number of patients with heart attacks, strokes, lung cancer, melanoma, symptomatic HIV, and osteosarcoma for each of the individual subpopulations as well. The effect of Craigslist introduction is insignificant for each condition in each of the subpopulations as well. Results are available upon request.

Table 13 **Exposure Model of Effect of Craigslist Introduction on Asymptomatic HIV (Negative Binomial)**

	Dependent variable							
	(1) *Total Count*	(2) *Total Caucasian*	(3) *Total African American*	(4) *Total Latino*	(5) *Total Non-Medicaid*	(6) *Total Medicaid*	(7) *Total Male*	(8) *Total Female*
Craigslist	0.127*** (0.0332)	0.0815 (0.0495)	0.146*** (0.0301)	0.155 (0.0982)	0.131*** (0.0290)	0.123** (0.0397)	0.128** (0.0466)	0.126*** (0.0261)
County Population	−2.81e−08 (6.59e−07)							
Caucasian Population		−1.83e−07 (8.55e−07)						
African American Population			1.25e−06 (1.29e−06)					
Latino Population				−6.80e−06** (2.27e−06)				
Population in Poverty					5.40e−07 (9.16e−07)			
Population Not in Poverty						−1.62e−07 (4.69e−07)		
Male Population							−7.86e−07 (1.33e−06)	
Female Population								8.50e−07 (1.69e−06)
Constant	4.034*** (0.521)	2.357*** (0.440)	3.511*** (0.309)	−0.0253 (0.116)	3.173*** (0.104)	3.511*** (0.324)	3.460*** (0.509)	3.106*** (0.685)
Hospital fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log pseudolikelihood	−9,085.12	−6,598.18	−6,623.45	−3,473.41	−5,819.06	−8,080.3	−7,621.05	−6,561.87

Notes. Number of observations = 4,349. Robust standard errors are in parentheses (clustered on county).
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

in addition to our hospital and time fixed effects. In effect, this measure captures the change in the population, as well as the individual subpopulations, over time. Prior to discussing results, we point out several potential complications with this analysis. First, because estimates of county population (both aggregate and by subpopulation) are not available at the quarter level, we use the annual population level for each estimate. Second, because estimates of the population of each county that is receiving Medicaid at a specific point in time are unavailable (this information is aggregated to the state-year level by the Centers for Medicare and Medicaid Services), we proxy the population of the county on Medicaid with the population of the county living below the poverty line. Because the primary exclusion criterion for *Medicaid* is income, this assumption is reasonable and the measures should be strongly correlated. The source of these data is the Area Resource File provided by the Health Resource and Services Administration (HRSA), and the Florida Department of Health.

Table 13 shows that results remain consistent in this analysis: there is a strong and significant effect of Craigslist entry on the asymptomatic HIV rate. Both genders are affected (men and women), and the effect exists across the socioeconomic spectrum (both Medicaid and non-Medicaid patients are affected, with the absolute effect on non-Medicaid being significantly larger (0.876371 versus 0.483935 patients per hospital quarter)). Moreover, the primary *racial* effect occurs within the African American community (with positive but insignificant coefficients within the Caucasian and Latino communities).

5.4. Exclusion of Untreated Hospitals

Whereas the exposure model suggests that the size of the relevant subpopulations is not driving the effect, it is possible that there is variation in the health of treated and untreated counties. Although our analysis excludes all hospitals who never treat an HIV patient, thereby plausibly eliminating communities where there is no HIV to spread, it is possible that untreated counties lack the necessary infection rate of HIV to experience an effect post Craigslist entry. Because this may introduce heterogeneity into treated and untreated groups, we next execute our estimations using only hospitals that eventually receive Craigslist treatment. In effect, this model allows us to compare the relative time trends in HIV between hospitals that receive Craigslist treatment earlier and those that receive treatment later. Results from this analysis, shown in Table 14, corroborate previous estimations and, interestingly, also show a significant effect within the Latino community.

Table 14 Negative Binomial Estimates of Total Number of Asymptomatic HIV Cases Using Only Treated Counties

	Dependent variable							
	(1) *Total Count*	(2) *Total Caucasian*	(3) *Total African American*	(4) *Total Latino*	(5) *Total Non-Medicaid*	(6) *Total Medicaid*	(7) *Total Male*	(8) *Total Female*
Craigslist	0.123***	0.0765	0.131***	0.230**	0.114**	0.135***	0.116**	0.139***
	(0.0332)	(0.0596)	(0.0282)	(0.0810)	(0.0398)	(0.0259)	(0.0423)	(0.0271)
Constant	3.999***	2.278***	3.758***	−0.182	3.386***	3.233***	3.158***	3.438***
	(0.0292)	(0.0451)	(0.0518)	(0.121)	(0.0495)	(0.0488)	(0.0527)	(0.0339)
Hospital fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log pseudolikelihood	−5,237.69	−3,679.29	−4,173.53	−2,393.83	−4,710.79	−3,539.87	−4,456.15	−3,930.49

Notes. Number of observations = 2,162. Robust standard errors are in parentheses (clustered on county).
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

5.5. Heterogeneity in Treated and Untreated Time Trends

Thus far findings are stable across multiple specifications and falsification tests: we see that neither changes in the population of treated counties nor bias in the selection of hospitals are influencing the main results. Nevertheless, it is still possible that there is unobserved heterogeneity in the time trends between treated and untreated counties that our previous analyses have not revealed. To mitigate this concern, we replicate our estimations using independent time fixed effects for treated and untreated counties. To the extent that this analysis allows for differences in the asymptomatic HIV trend across the treated and untreated groups, it should reinforce the claim that the hospital fixed effects have effectively controlled for ex ante heterogeneity in the groups. Results are available in Table 15 and remain consistent with the largest absolute effect accruing to non-Medicaid patients and African Americans. Furthermore, consistent with the relative time model, we see a small increase of 0.29 patients per hospital quarter in the Latino community.

5.6. Change in Testing Behavior

Thus far our empirical strategy in the robustness checks has focused on eliminating unobserved heterogeneity between the treated and untreated groups. However, one further plausible alternative explanation for the change in the asymptomatic HIV incidence rate, post-Craigslist implementation, is that there is a change in the propensity for individuals to get tested for sexually transmitted diseases. For example, it is possible that individuals who habitually engage in risky behavior are using the website to solicit casual partners who are outside their standard pool of casual partners. If this is the case, they may worry about the possibility of contracting disease and choose to get tested for STDs after implementation at a higher rate, thereby increasing the number of HIV diagnoses without actually influencing the underlying incidence rate. To mitigate this alternative explanation, we use data from the Behavioral Risk Factor Surveillance System (BRFSS) survey conducted by the CDC.[32]

The quinquennial BRFSS survey contains county-level information on the percentage of the population that reports having been tested for HIV in the previous year. We compare the propensity for individuals to be tested for HIV in 2002 and 2007 (the first year of our sample and the year immediately following our sample period) across the regions represented in our data. Results are available in Table 16. *Craigslist Treated* indicates counties that receive the Craigslist treatment during the course of our study, and *Untreated* indicates counties that do not. We find no significant difference in the ex ante or ex post propensity for individuals to get tested in treated or untreated counties. Furthermore, there is no significant difference in the change in propensity to get tested for HIV between 2002 and 2007. These results further affirm the assertion that it is Craigslist that is causing the observed increase in HIV.[33]

5.7. Entry Model

One further concern that is present in our data is exogeneity of the entry decision by Craigslist. Although our models, thus far, have indicated that there is no heterogeneity in the pretreatment asymptomatic HIV trend, it is possible that there is stable variation in the number of HIV patients across treated and untreated hospitals. To ensure that the entry of Craigslist is not systematically related to the incidence of HIV within the local area, we estimate an entry model to test whether the number of HIV patients in a local area is positively correlated with entry. To the extent that a larger population of HIV-positive carriers will magnify the negative effect of

[32] See http://www.cdc.gov/brfss/ (accessed November 20, 2014).

[33] It is important to note that, because of the small sample size, these analyses should be interpreted with caution. We thank the anonymous reviewer for this point.

Table 15 Negative Binomial Estimates of Total Number of Asymptomatic HIV Cases Using Independent Time Fixed Effects for Hospitals That Receive Treatment and Those That Never Receive Treatment

	Dependent variable							
	(1) *Total Count*	(2) *Total Caucasian*	(3) *Total African American*	(4) *Total Latino*	(5) *Total Non-Medicaid*	(6) *Total Medicaid*	(7) *Total Male*	(8) *Total Female*
Craigslist	0.122***	0.0742	0.133***	0.230**	0.112**	0.138***	0.116**	0.140***
	(0.0328)	(0.0574)	(0.0272)	(0.0805)	(0.0390)	(0.0263)	(0.0418)	(0.0277)
Constant	4.276***	2.510***	4.029***	−0.0102	3.707***	3.347***	3.423***	3.653***
	(0.0668)	(0.0840)	(0.0633)	(0.0744)	(0.0703)	(0.0512)	(0.0664)	(0.0452)
Hospital fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Treated quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Untreated quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log pseudolikelihood	−9,075.79	−6,589.29	−6,615.55	−3,461.24	−8,069.23	−5,812.54	−7,611.61	−6,555.67

Notes. Number of observations = 2,452. Robust standard errors are in parentheses (clustered on county).
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Craigslist entry (by increasing the likelihood of an encounter with an HIV carrier), it is important to rule out this threat. To conduct this analysis, we execute a logit hazard specification (Singer and Willett 1993) with Craigslist entry as the dependent variable. Independent variables include a broad range of local socioeconomic conditions such as population, age, income, poverty level, and education, as well as a number of HIV cases diagnosed since 1990 (the earliest year of data availability), in the aggregate and for each subpopulation examined in the main model. As with the exposure model, the additional data required for these estimations are retrieved from the Area Resource File. The unit of analysis remains the hospital quarter. Results are available in Table 17.

As would be expected, we see that an increase in the number of college graduates in a county is correlated with an increased likelihood of Craigslist entry. Furthermore, an increase in the number of younger users (aged 20–44) is correlated with an increased likelihood of entry. However, results indicate that the number of previously diagnosed HIV cases at the focal hospital, both in the aggregate and in the individual subpopulations, is not positively influencing the likelihood of entry by Craigslist.

Table 16 Two-Sample t-Test of Population Propensity to Be Tested for HIV Before and After Treatment (t-Value Indicates the Outcome of the Two-Sample t-Test Comparing Treated and Untreated Counties)

	Period		
	2002	2007	Δ from 2002–2007
Craigslist Treated			
μ	21.3583333	21.275	−0.083
σ	5.51979221	5.77252505	2.2993
N	12	12	12
Untreated			
μ	20.6714286	19.0303571	−1.641
σ	4.82462299	6.22893307	6.2397
N	56	56	56
t-Value	0.4015	1.2083	1.46

5.8. Random Implementation Model

As a final robustness check, we estimate a random implementation model. Although our results have been consistent across a wide variety of specifications, it is still possible that the increase in the HIV infection rate is related to an idiosyncrasy associated with Craigslist treated hospitals, and not as a result of Craigslist implementation. If this were true, the effect would be significant with any ordering of Craigslist introduction in the hospitals that were eventually treated. We conduct further analysis to determine how likely it is that a *random* implementation of Craigslist would yield an aggregate effect size that is comparable to our estimates.[34] We estimate the probability of randomly finding the aggregate effect in three ways. First, we randomly treat 843 hospital quarters (the number of treated hospital quarters in our data) and regress the total number of HIV patients, by hospital quarter, on this dummy variable (*Pseudo-Treated*) along with hospital and time fixed effects. The coefficient of the estimates for *Pseudo-Treated* is then stored. After replicating this random treatment 500 times, we calculate the mean and standard deviation of the *Pseudo-Treated* coefficients. From this we calculate the Z-score of the difference between our estimated coefficient (our original β estimate) and the mean of the randomly calculated coefficients.

After completing the purely random implementation model, we replicate the procedure treating only the hospitals that eventually receive *Craigslist*. Finally, we randomly treat hospitals based on the time they receive the Craigslist treatment. To conduct this

[34] We thank an anonymous reviewer for this suggestion.

Table 17 **Outcome of Logit Hazard Model for the Entry of Craigslist Diagnosed Cases to Date Indicating the Number of Asymptomatic Cases by Subpopulation Diagnosed at Hospital j Since 1990 (First Date of Data Availability) (Dependent Variable = *Craigslist*)**

	(1)	(2)	(3)	(4)	(5)	(6)
Population	−4.61e−06	−4.48e−06	−5.42e−06	−4.55e−06	−4.43e−06	−5.44e−06
	(5.96e−06)	(5.96e−06)	(6.04e−06)	(5.97e−06)	(5.96e−06)	(6.04e−06)
African American Population	−1.64e−05**	−1.70e−05**	−1.78e−05**	−1.70e−05**	−1.68e−05**	−1.85e−05**
	(6.07e−06)	(6.14e−06)	(6.31e−06)	(6.14e−06)	(6.12e−06)	(6.52e−06)
Population in Poverty	−8.75e−06	−8.35e−06	−9.15e−06	−8.24e−06	−8.49e−06	−9.04e−06
	(6.78e−06)	(6.82e−06)	(6.95e−06)	(6.83e−06)	(6.81e−06)	(7.01e−06)
Population Age 20–44 (0000s)	0.455***	0.456***	0.472***	0.458***	0.454***	0.476***
	(0.128)	(0.128)	(0.131)	(0.128)	(0.128)	(0.132)
Per-Capita Income	0.884**	0.878*	0.915**	0.882*	0.877*	0.916**
	(0.342)	(0.343)	(0.346)	(0.343)	(0.343)	(0.345)
Population w/High School Diploma (0000s)	−0.287***	−0.290***	−0.266**	−0.289***	−0.290***	−0.267**
	(0.0858)	(0.0863)	(0.0867)	(0.0864)	(0.0863)	(0.0864)
Population w/College Diploma (0000s)	0.622*	0.621*	0.598*	0.618*	0.624*	0.600*
	(0.253)	(0.253)	(0.256)	(0.253)	(0.253)	(0.254)
Diagnosed HIV Cases to Date		0.000197				
		(0.000204)				
Diagnosed Caucasian HIV Cases to Date			−0.00443			−0.00265
			(0.00245)			(0.00700)
Diagnosed AA HIV Cases to Date			0.00166			0.00396
			(0.000919)			(0.00537)
Diagnosed Latino HIV Cases to Date			0.00575			0.00813
			(0.00492)			(0.00830)
Diagnosed Medicaid HIV Cases to Date				0.000563		−0.00299
				(0.000528)		(0.00622)
Diagnosed Male HIV Cases to Date					0.000295	−0.00177
					(0.000380)	(0.00620)
Constant	−6.908***	−6.912***	−6.932***	−6.923***	−6.905***	−6.955***
	(1.302)	(1.302)	(1.309)	(1.303)	(1.302)	(1.309)
Log pseudolikelihood	−171.07	−170.64	−168.54	−170.53	−170.79	−168.41

Notes. Number of observations = 1,067. Robust standard errors are in parentheses. AA, African American.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

final analysis, we change the time of implementation between the hospitals in different counties at random. For example, in this analysis a hospital in Tampa (originally treated in November 2003) might instead receive a treatment date of October 2002 (the time of treatment in Miami). The hospital in Miami then receives the treatment time of a hospital in Orlando (February 2004). As with our previous random implementation models, these models are replicated 500 times. Results are available in Table 18

Table 18 **Outcome of Random Implementation Model**

	Random implementation	Random implementation in treated	Random implementation of treatment time
μ of random β	0.00004	−0.00050	0.00030
σ random β	0.01819	0.01647	0.05020
Estimated β	0.12600	0.12600	0.12600
Replications	500	500	500
Z-score	6.92599	7.67856	2.50423
p-Value	< 0.0001	< 0.0001	< 0.01

and indicate, in each of the three models, that our estimated effect size is significantly larger than what would be expected purely by chance. Because our randomly determined treatment β is statistically indistinguishable from zero, this further supports the conclusion that there is no significant unaccounted for difference between treated and untreated hospitals, thereby strengthening our causal claim.

6. Implications and Conclusion

Our work was motivated by the growing role of the Internet in individual and public health and the documented variation in Internet use and digital capabilities across sociodemographic segments of society. Whereas Internet-enabled two-sided matching platforms offer robust benefits in terms of facilitating social contact, they may also yield connections that pose risks. We examined the effect of a platform for the solicitation of sexual partners on the local incidence rate of asymptomatic HIV. We asked, to whom do the negative effects of platform entry accrue based on the sociodemographic characteristics of race, gender, and socioeconomic status? Although the documented HIV incidence rate is highest among ethnic minorities and the socioeconomic lower class (CDC 2011), there is variation in the nature and extent of Internet usage among these groups (Barry 2013, DiMaggio et al. 2004, Hargittai 2010, Warschauer 2004). Using available data, we developed baseline estimates of the subpopulations who are likely to be HIV carriers and also active on the site to gain ex ante insights into where the negative externalities may be experienced. Recognizing that the assumptions of the model limited its generalizability, we further conducted a detailed empirical analysis of the effect of platform entry.

Analysis of a census of approximately 12 million patients who are subjected to a natural experiment yields five main results. First, consistent with the state-level effect found in Chan and Ghose (2014), the entry of Craigslist significantly increases the asymptomatic HIV incidence rate for residents of treated areas (both the focal cities and the surrounding suburban areas). Second, the absolute and relative increase in the HIV incidence rate is significantly larger among one historically at-risk population, African Americans. Third, across SESs, the absolute increase is highest in one historically less at-risk population, those of medium and high SES. Fourth, we find that men and women do not experience significantly different penalties, even though men are at higher risk for HIV and women are modestly less likely to use the Internet. Finally, the effects on Caucasians and Latinos are statistically indistinguishable from each other despite the fact that Latinos are at a greater risk of HIV infection.

We compute the economic implications of sites that facilitate such behavior. Annually, results indicate an increased cost of \$6.896 million per year as of 2006 in the state of Florida.[35] In the aggregate, results suggest that roughly 1,149 patients have been admitted to Florida hospitals who otherwise would not have been admitted during the five-year period of our study (62.93% of them African American[36]). At an average cost of \$618,000 over the patient's lifetime (Schackman et al. 2006), this translates to an additional financial burden of \$710 million in the state of Florida alone. We note that these estimates are conservative because the diagnosis rate during the asymptomatic phase of the disease is not 100% (Janssen et al. 1992). When compared with the estimates reported in Chan and Ghose (2014), ours are broadly consistent (a 13.5% rise relative to their 15.9%).

This study further contributes to inquiry into digital inequality and variations in digital capabilities across subpopulations (DiMaggio et al. 2004, Warschauer 2004), research that has historically focused on the negative downstream implications of decreased information access and connectivity. Even within the health policy literature, discussions have been confined to the undesirable effect of limited Internet access on the dissemination of health information (Brodie et al. 2000). Our results suggest that although lower access to online resources among traditionally at-risk groups has been acknowledged, penalties continue to accrue to them disproportionately. Insofar as African American subpopulations have also been widely documented as experiencing health disparities (CDC 2005), we find that Craigslist exacerbates these differences. Whereas limited access to the resources of the Internet may reduce social welfare in most instances, we describe a setting where use of an online platform diminishes welfare further. For policy makers, this emphasizes the need for greater vigilance regarding the unintended consequences of online marketplaces that may disproportionately be borne by select subpopulations. Other examples of such consequences are not difficult to find: a recent study of Airbnb, a site that facilitates matching of buyers and sellers for rental properties, finds that African American hosts charge significantly lower prices than Caucasian hosts for equivalent properties (Edelman and Luca 2014). Although access to platforms cannot be limited through policy, enhanced awareness of these consequences may allow policy makers to take countermeasures, such as public information campaigns, and also motivate platform owners to take steps such as a conspicuous display of the potential for increased risk, to mitigate adverse outcomes.

Relatedly, this study highlights the importance of further investigation into when the beneficial and punitive effects of increased Internet access are experienced for populations across the spectrum of sociodemographic characteristics. Our finding that even populations considered to be at lower risk for HIV infection (such as women and Caucasians) experience an increase in HIV incidence underscores the need for broader communication and dissemination of the risks posed by the type of online matching platforms studied here. It is possible that these populations worry less about contracting the disease because they have historically had lower risk and, as a result, are not sensitive to the increased threat posed by platform use. Once again, public information campaigns

[35] This calculation is derived from the 125 treated hospitals at the conclusion of the sample using the aggregate increase of 1.136268 patients per treated hospital quarter and an annual cost of \$10,121.60 (Chesson et al. 2004).

[36] Impact on the subpopulation is calculated by dividing the African American marginal effect (0.857629 patients per hospital quarter) by the aggregate marginal effect.

may be necessary to communicate the risk of platform use more broadly rather than targeted only at high-risk populations.[37]

Although results from the empirical estimations do not allow us to clearly isolate the mechanism by which the increase in HIV is occurring, they nevertheless offer insights into the potential mechanisms more and less likely to be at play. First, the relative parity, i.e., the lack of a statistically significant difference, between men and women indicates that the increase is likely not simply a result of increased unprotected homosexual sex between men, as previously suggested in other studies (Benotsch et al. 2002, Rosser et al. 2009) and recent media reports (see Holley 2015). Second, the congruence between the baseline numerical estimates and the empirical results suggests that platforms such as Craigslist are not increasing the rate of out-group dating by race, insofar as African Americans experience the largest effect, with significantly smaller effects for Caucasians and Latinos.

Our results also point to two plausible mechanisms that may underlie the observed effects. It may be the case that Craigslist is increasing the density of sexual networks by filling structural holes that exist between infected and uninfected networks of sexual partners. Alternatively, the anonymity of Craigslist may be increasing the amount of latent homosexual dating by men living publicly as heterosexuals. As this would result in the virus also being passed to female partners of the infected male, it may explain the lack of significant difference in the effect between men and women. Insofar as either of these mechanisms is plausible, the ambiguity in the dominance of either effect offers useful opportunities for future research.

When considering the effect on individual subpopulations, our study raises an additional set of questions worthy of further exploration. First, given the relative digital disadvantage among African Americans, why is the incidence rate increase for this group so large? One plausible explanation for this effect is the common misconception about the digital divide that it is binary (Warschauer 2004), with individuals either having or not having access to online resources. In contrast to this view, digital disparity is more often a continuum ranging from no access, to usage of public access points (such as libraries during the time of the study; see Heuertz et al. 2002, Moore et al. 2002), to unskilled and then skilled exploitation of online resources in the home (DiMaggio et al. 2004, Hargittai 2010). Although research suggests that the digitally disadvantaged rarely utilize online resources for welfare-enhancing activities (Hargittai 2010, Zillien and Hargittai 2009), this does not imply that there is no utilization.

A second interesting question raised by our results relates to the small and intermittent effect observed in the Latino population. Recall that although our main estimation did not show a significant effect for Latinos, we found an effect in several of the other estimations and robustness checks. To the degree that Latinos are considered to be a disadvantaged population that is disproportionately affected by HIV (CDC 2011), this begs the question why the Latino effect is qualitatively similar to the effect on Caucasians. Although our methodology does not allow us to determine the exact mechanism, several possibilities exist. The first is language barriers (Barry 2013). Because 34.8 million Hispanic households (65% of the population) speak Spanish as the primary language in the home (Gonzalez-Barrera and Lopez 2013), and the American Craigslist interface is written in English, it is possible that language constraints prevent widespread use of the platform among the Latino community. A second possibility is limited detection and seeking of treatment. Florida is home to more than one million undocumented Latinos during the time of our investigation (see Passel and Cohn 2011), so it may be the case that these individuals are not seeking treatment out of fear of deportation. Finally, it is plausible that the Latino community in Florida is not representative of the Latino population nationally (as evidenced by the significantly higher population proportion and its relative wealth, as indicated by the firm ownership Latinos enjoy in Florida[38]). As a consequence, it is possible that the differing socioeconomic status this subpopulation has within the Florida community has yielded different results.

Our finding that the effect for men and women is not statistically different is also notable. Studies on gender differences in Internet use suggest that men are more likely to use the Internet (Hargittai 2010), although this gap is narrowing with the increased availability and use of technology among youth. Research also suggests that women tend to use the Internet to a greater extent for social relationships than instrumental transactions, such as banking, compared with men (Weiser 2000). The fact that both genders significantly increase their probability of contracting HIV upon the availability of a matching platform for casual encounters is another instance of a potentially

[37] Although there are general information campaigns for HIV (targeting the population as a whole), the majority of ongoing campaigns from the CDC (e.g., Act Against AIDS) and AIDS.gov (e.g., HIV/AIDS Awareness Days) target "at-risk" subpopulations (e.g., African Americans, Latinos, bisexual and homosexual men) and do not directly address the danger faced by either non-African American women or those of relatively higher SES.

[38] See http://quickfacts.census.gov/qfd/states/12000.html (accessed February 20, 2015).

undesirable "equalizing" effect of the Internet. Alternatively, this finding may also suggest that bisexual and homosexual men are engaging in safe sex more often when leveraging online platforms for the solicitation of casual partners (thereby decreasing the effect for men to a level where it is statistically indistinguishable from that for women). Insofar as these evolving sexual dynamics may offer a host of insights for public policy and public health researchers, this remains a fruitful avenue for future research.

We acknowledge the limitations of this study. First, we cannot observe the Craigslist utilization rates for visitors to treated areas (from either inside or outside the state of Florida). Econometrically, this is of limited concern because it will bias results downward, making estimates more conservative (because these visitors will likely be diagnosed at the untreated hospitals near their homes, thereby introducing parity into the increase in the HIV rate across the two groups). However, from the perspective of epidemiological research, further work investigating the effect of platforms that facilitate the transmission of disease across geographical regions is needed. Second, our data do not allow us to observe patient readmittance. Although Florida is experiencing a decreasing HIV incidence rate during the study period, and advances in medical treatment are causing patient readmittance to slow, this is clearly a data limitation. Third, the use of Medicaid as a proxy for low SES can be improved upon. However, we note that because of privacy and confidentiality concerns, detailed income data at the individual level that is tied to medical records are difficult to obtain. Fourth, as Craigslist offers many opportunities for couples to meet and interact, our results cannot ensure that the increase in HIV incidence is exclusively a result of casual sexual encounters. It is plausible, for example, that long-term dating enabled by the site also plays a role in the increase in HIV prevalence or that individuals who engage in casual sex contract the disease and then pass it on to their spouses.

Further, our data do not allow us to perfectly track population migration over time, because the composition of the population within subpopulations may be changing. Although our empirical results from the exposure model suggest that changes in each of the subpopulation levels are not influencing the effect and there is no concomitant spike in other conditions, we are unable to completely rule this explanation out. Our data also do not allow us to observe the sociodemographic characteristics of the individual who passed the HIV virus to the patient. One future, important extension of this work will be to investigate how the focal individual's selection of sexual partner on matching platforms influences the spread of disease. We are unable to capture the effect on patients who do not seek treatment. However, in the absence of changes in testing behavior between treated and untreated counties, this should not bias our results because the selection into seeking treatment does not appear to be correlated with Craigslist entry. Econometrically, as discussed by Allison and Waterman (2002), there are concerns with the present encoding of the negative binomial regression relating to overdispersion. Although we have used several other estimators to corroborate the results, this is a fundamental limitation of the estimation procedure. Finally, although the economic implications of the HIV externality are large, it is important to note that we have made no attempt in this work to quantify the positive effects of Craigslist. Because significant public welfare may be generated by these platforms, it would be inappropriate to draw any conclusions regarding the net effect of Craigslist as a whole, whether positive or negative, on public welfare from our results.

Acknowledgments

The authors thank the department editor, the associate editor, and the anonymous reviewers for the many helpful suggestions that significantly improved the quality of this work. The paper has benefited from feedback received from participants and discussants at the Workshop on Information Systems Economics, the Workshop on Health IT and Economics, University of Pennsylvania, University of Arizona, University of Minnesota, University of Virginia, New York University, and McGill University. The authors also thank Gordon Burtch, Seth Carnahan, Jason Chan, Daniel Malter, Jeffrey McCullough, Mahka Moeen, Dan Olson, and Kislaya Prasad for their comments on earlier versions. All remaining errors are the authors' own.

References

Agarwal R, Gao G, DesRoches C, Jha AK (2010) The digital transformation of healthcare: Current status and the road ahead. *Inform. Systems Res.* 21(4):796–809.

Ai C, Norton EC (2003) Interaction terms in logit and probit models. *Econom. Lett.* 80(1):123–129.

Allison PD, Waterman RP (2002) Fixed-effects negative binomial regression models. *Sociol. Methodol.* 32(1):247–265.

American Psychological Association (2013) Socioeconomic status. Accessed August 8, 2013, http://www.apa.org/topics/socioeconomic-status/.

Angrist JD, Pischke J-S (2008) *Mostly Harmless Econometrics: An Empiricist's Companion* (Princeton University Press, Princeton, NJ).

Bakos Y, Bailey J (1997) An exploratory study of the emerging role of electronic intermediaries. *Internat. J. Electronic Commerce* 1(3):7–20.

Bapna R, Ramaprasad J, Shmueli G, Umyarov A (2012) One-way mirrors in online dating: A randomized field experiment. *Workshop Inform. Systems Econom., Orlando, FL.*

Barry JJ (2013) Crossing the digital divide: Is access to the Internet an economic right? Working paper, University of Connecticut, Storrs. http://dx.doi.org/10.2139/ssrn.2315307.

Benotsch EG, Kalichman S, Cage M (2002) Men who have met sex partners via the Internet: Prevalence, predictors, and implications for HIV prevention. *Arch. Sexual Behav.* 31(2):177–183.

Bingham TA, Harawa NT, Johnson DF, Secura GM, MacKellar DA, Valleroy LA (2003) The effect of partner characteristics on HIV infection among African American men who have sex with men in the young men's survey, Los Angeles, 1999–2000. *AIDS Ed. Prevention* 15(1, Supplement):39–52.

Bozzette SA, Berry SH, Duan N, Frankel MR, Leibowitz AA, Lefkowitz D, Emmons C-A, Senterfitt JW, Berk ML, Morton SC (1998) The care of HIV-infected adults in the United States. *New England J. Medicine* 339(26):1897–1904.

Brodie M, Flournoy RE, Altman DE, Blendon RJ, Benson JM, Rosenbaum MD (2000) Health information, the Internet, and the digital divide. *Health Affairs* 19(6):255–265.

Brown P (1988) *The Body and Society: Men, Women, and Sexual Renunciation in Early Christianity* (Columbia University Press, New York).

Brynjolfsson E, Smith MD (2000) Frictionless commerce? A comparison of Iinternet and conventional retailers. *Management Sci.* 46(4):563–585.

Burke MA, Fournier GM, Prasad K (2003) Physician social networks and geographical variation in medical care. CSED Working Paper 33, Center on Social and Economic Dynamics, Brookings Institution, Washington, DC.

Burke MA, Fournier GM, Prasad K (2007) The diffusion of a medical innovation: Is success in the stars? *Southern Econom. J.* 73(3):588–603.

Campsmith ML, Rhodes P, Hall HI, Green T (2008) HIV prevalence estimates—United States, 2006. *Morbidity Mortality Weekly Rep.* 57(39):1073–1076.

CDC (2005) Health disparities experienced by black or African Americans—United States. *Morbidity Mortality Weekly Rep.* 54(1):1–3.

CDC (2008) Revised surveillance case definitions for HIV infection among adults, adolescents, and children aged < 18 months and for HIV infection and AIDS among children aged 18 months to < 13 years—United States, 2008. *Morbidity Mortality Weekly Rep.* 57(RR-10):1–8.

CDC (2011) HIV surveillance—United States, 1981–2008. *Morbidity Mortality Weekly Rep.* 60(21):689–693.

CDC (2013) Deaths and mortality. Accessed February 20, 2015, http://www.cdc.gov/nchs/fastats/deaths.htm.

Chan J, Ghose A (2014) Internet's dirty secret: Assessing the impact of online intermediaries on the outbreak of sexually transmitted diseases. *MIS Quart.* 38(4):955–976.

Chesson HW, Blandford JM, Gift TL, Tao G, Irwin KL (2004) The estimated direct medical cost of sexually transmitted diseases among American youth, 2000. *Perspect. Sexual Reproductive Health* 36(1):11–19.

DiMaggio P, Hargittai E, Celeste C, Shafer S (2004) Digital inequality: From unequal access to differentiated use. Neckerman K, ed. *Social Inequality* (Russell Sage Foundation, New York), 355–400.

Dorland (2011) *Dorland's Illustrated Medical Dictionary*, 32nd ed. (Elsevier–Health Sciences Division, Philadelphia).

Dugas AF, Hsieh Y-H, Levin SR, Pines JM, Mareiniss DP, Mohareb A, Gaydos CA, Perl TM, Rothman RE (2012) Google Flu Trends: Correlation with emergency department influenza rates and crowding metrics. *Clinical Infectious Diseases* 54(4):463–469.

Edelman B, Luca M (2014) Digital discrimination: The case of Airbnb.com. HBS NOM Unit Working Paper 14-054, Harvard Business School, Boston.

Elford J, Bolding G, Sherr L (2001) Seeking sex on the Internet and sexual risk behaviour among gay men using London gyms. *AIDS* 15(11):1409–1415.

Fallows D (2005) How women and men use the Internet. Report, Pew Research Center, Washington, DC. http://www.pewinternet.org/2005/12/28/how-women-and-men-use-the-internet/.

Finzi D, Hermankova M, Pierson T, Carruth LM, Buck C, Chaisson RE, Quinn TC, et al. (1997) Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science* 278(5341):1295–1300.

Floyd FJ, Stein TS (2002) Sexual orientation identity formation among gay, lesbian, and bisexual youths: Multiple patterns of milestone experiences. *J. Res. Adolescence* 12(2):167–191.

Fox S (2005) Digital divisions: There are clear divisions among those with broadband connections, dial-up connections, and no connections at all to the Internet. Pew Internet and American Life Project report, Pew Research Center, Washington, DC. http://www.pewinternet.org/files/old-media/Files/Reports/2005/PIP_Digital_Divisions_Oct_5_2005.pdf.pdf.

Fox S, Livingston G (2007) III. Latinos online. Report, Pew Research Center, Washington, DC. http://www.pewhispanic.org/2007/03/14/iii-latinos-online/.

Garofalo R, Herrick A, Mustanski BS, Donenberg GR (2007) Tip of the iceberg: Young men who have sex with men, the Internet, and HIV risk. *Amer. J. Public Health* 97(6):1113–1117.

Goh JM, Gao G, Agarwal R (2016) The creation of social value: Can an online health community reduce rural–urban health disparities? *MIS Quart.* 40(1):247–263.

Gonzalez-Barrera A, Lopez H (2013) Spanish is the most spoken non-English language in U.S. homes, even among non-Hispanics. Report, Pew Research Center, Washington, DC. http://www.pewresearch.org/fact-tank/2013/08/13/spanish-is-the-most-spoken-non-english-language-in-u-s-homes-even-among-non-hispanics/.

Greenwood BN, Agarwal R, Agarwal R, Gopal A (2013) A tale of three tensions: Changes in decision making after information shocks. *Acad. Management Proc.* 2013(Meeting abstract supplement):Article 16890.

Grov C (2010) Risky sex-and drug-seeking in a probability sample of men-for-men online bulletin board postings. *AIDS Behav.* 14(6):1387–1392.

Hargittai E (2010) Digital na(t)ives? variation in Internet skills and uses among members of the "net generation." *Sociol. Inquiry* 80(1):92–113.

Heinzl H, Waldhör T, Mittlböck M (2005) Careful use of pseudo *R*-squared measures in epidemiological studies. *Statist. Medicine* 24(18):2867–2872.

Heuertz L, Gordon A, Moore E, Gordon M (2002) *Public Libraries and the Digital Divide: How Libraries Help* (University of Washington, Seattle).

Hitsch GJ, Hortaçsu A, Ariely D (2010) Matching and sorting in online dating. *Amer. Econom. Rev.* 100(1):130–163.

Hoetker G (2007) The use of logit and probit models in strategic management research: Critical issues. *Strategic Management J.* 28(4):331–343.

Holley P (2015) When Craigslist comes to town, HIV infections go up, study says. *Washington Post* (February 2), http://www.washingtonpost.com/news/morning-mix/wp/2015/02/02/when-craigslist-comes-to-town-hiv-infections-go-up-study-says/.

Holmberg SD, Moorman AC, Greenberg AE, Committee DS (2004) Trends in rates of myocardial infarction among patients with HIV. *New England J. Medicine* 350(7):730–732.

Hsieh JJP-A, Rai A, Keil M (2011) Addressing digital inequality for the socioeconomically disadvantaged through government initiatives: Forms of capital that affect ICT utilization. *Inform. Systems Res.* 22(2):233–253.

Infoplease (2015) QuickFacts from the US Census Bureau: Florida. Accessed February 20, 2015, http://www.washingtonpost.com/news/morning-mix/wp/2015/02/02/when-craigslist-comes-to-town-hiv-infections-go-up-study-says/.

Janssen RS, St. Louis ME, Satten GA, Critchley SE, Petersen LR, Stafford RS, Ward JW, et al. (1992) HIV infection among patients in U.S. acute care hospitals. *New England J. Medicine* 327(7):445–452.

Jones S (2009) Online classifieds: The number of online adults to use classified ads websites, such as Craigslist, more than doubled from 2005 to 2009. Pew Internet and American Life Project report, Pew Research Center, Washington, DC. http://www.pewinternet.org/~/media//Files/Reports/2009/PIP%20-%20Online%20Classifieds.pdf.

Kim AA, Kent C, McFarland W, Klausner JD (2001) Cruising on the Internet highway. *J. Acquired Immune Deficiency Syndromes* 28(1):89–93.

King G, Zeng L (2001) Explaining rare events in international relations. *Internat. Organ.* 55(3):693–715.

Little SJ, Holte S, Routy J-P, Daar ES, Markowitz M, Collier AC, Koup RA, Mellors JW, Connick E, Conway B (2002) Antiretroviral-drug resistance among patients recently infected with HIV. *New England J. Medicine* 347(6):385–394.

MacDonald K (1995) The establishment and maintenance of socially imposed monogamy in Western Europe. *Politics Life Sci.* 14(1):3–23.

Malone T, Yates J, Benjamin R (1987) Electronic markets and electronic hierarchies. *Comm. ACM* 30(6):484–497.

Miguelino-Keasling V (2010) Poverty and HIV infection: 2006–2007 NHBS national and San Diego findings. Report, National HIV Behavioral Surveillance, Centers for Disease Control, Atlanta.

Moore E, Gordon AC, Gordon MT, Heuertz L (2002) It's working: People from low-income families disproportionately use library computers. Report to the Bill & Melinda Gates Foundation, Evans School of Public Affairs, University of Washington, Seattle.

National Center for Health Statistics (1998) *International Classification of Diseases, Ninth Revision (ICD-9)*. Accessed October 28, 2015, http://www.cdc.gov/nchs/icd/icd9.htm.

National Telecommunications and Information Administration/Economics and Statistics Administration (NTIA/ESA) (2013) Exploring the digital nation: America's emerging online experience. Report, U.S. Department of Commerce, Washington, DC.

Office of Economic and Demographic Research (2007) Florida demographic overview. Report, Florida Legislature, Tallahassee.

Ottaviani G, Jaffe N (2010) The epidemiology of osteosarcoma. Jaffe N, Bruland ØS, Bielack SS, eds. *Pediatric and Adolescent Osteosarcoma* (Springer, New York), 3–13.

Padgett PM (2007) Personal safety and sexual safety for women using online personal ads. *Sexuality Res. Soc. Policy* 4(2):27–37.

Palella FJ Jr, Baker RK, Moorman AC, Chmiel JS, Wood KC, Brooks JT, Holmberg SD, HIV Outpatient Study Investigators (2006) Mortality in the highly active antiretroviral therapy era: Changing causes of death and disease in the HIV Outpatient Study. *J. Acquired Immune Deficiency Syndromes* 43(1):27–34.

Panel on Antiretroviral Guidelines for Adults and Adolescents (2006) Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. Report, Department of Health and Human Services, Washington, DC. https://aidsinfo.nih.gov/contentfiles/lvguidelines/AdultAndAdolescentGL.pdf.

Pantaleo G, Graziosi C, Fauci AS (1993) The immunopathogenesis of human immunodeficiency virus infection. *New England J. Medicine* 328(5):327–335.

Parsons JT, Tesoriero JM, Carballo-Dieguez A, Hirshfield S, Remien RH (2006) HIV behavioral research online. *J. Urban Health* 83(1):73–85.

Passel JS, Cohn D (2011) Unauthorized immigrant population: National and state trends, 2010. Report, Pew Research Center, Washington, DC. http://www.pewhispanic.org/2011/02/01/unauthorized-immigrant-population-brnational-and-state-trends-2010/.

Quenqua D (2009) Recklessly seeking sex on the Internet. *New York Times* (April 17), http://www.nytimes.com/2009/04/19/fashion/19craigslist.html?_r=0.

Rogers MK (2010) The psyche of cybercriminals: A psycho-social perspective. Ghosh S, Turrini E, eds. *Cybercrimes: A Multidisciplinary Analysis* (Springer, Berlin), 217–235.

Rosenfeld MJ, Thomas RJ, Falcon M (2011) How Couples Meet and Stay Together, waves 1, 2, and 3: Public version 3.04. (Data set.) Stanford University Libraries, Stanford, CA.

Rosenfeld MJ, Thomas RJ, Falcon M (2014) How Couples Meet and Stay Together, waves 1, 2, and 3: Public version 3.04, plus wave 4 supplement version 1.02 and wave 5 supplement version 1.0. (Data set.) Stanford University Libraries, Stanford, CA.

Rosser BR, Oakes JM, Horvath K, Konstan J, Danilenko G, Peterson J (2009) HIV sexual risk behavior by men who use the Internet to seek sex with men: Results of the Men's INTernet Sex Study-II (MINTS-II). *AIDS Behav.* 13(3):488–498.

Schackman BR, Gebo KA, Walensky RP, Losina E, Muccio T, Sax PE, Weinstein MC, Seage GR III, Moore RD, Freedberg KA (2006) The lifetime cost of current HIV care in the United States. *Medical Care* 44(11):990–997.

Seamans R, Zhu F (2014) Responses to entry in multi-sided markets: The impact of Craigslist on local newspapers. *Management Sci.* 60(2):476–493.

Secretary's Advisory Committee on National Health Promotion and Disease Prevention Objectives for 2020 (2008) Recommendations for the framework and format of Healthy People 2020. Report, U.S. Department of Health and Human Services, Washington, DC.

Singer JD, Willett JB (1993) It's about time: Using discrete-time survival analysis to study duration and the timing of events. *J. Ed. Behav. Statist.* 18(2):155–195.

Warschauer M (2004) *Technology and Social Inclusion: Rethinking the Digital Divide* (MIT Press, Cambridge, MA).

Weiser EB (2000) Gender differences in Internet use patterns and Internet application preferences: A two-sample comparison. *CyberPsych. Behav.* 3(2):167–178.

Williamson OE (1981) The economics of organization: The transaction cost approach. *Amer. J. Sociol.* 87(3):548–577.

Zelner BA (2009) Using simulation to interpret results from logit, probit, and other nonlinear models. *Strategic Management J.* 30(12):1335–1348.

Zenilman JM, Ellish N, Fresia A, Glass G (1999) The geography of sexual partnerships in Baltimore: Applications of core theory dynamics using a geographic information system. *Sexually Transmitted Diseases* 26(2):75–81.

Zickuhr K, Smith A (2012) Digital differences. Pew Internet and American Life Project report, Pew Research Center, Washington, DC. http://www.pewinternet.org/files/old-media/Files/Reports/2012/PIP_Digital_differences_041312.pdf.

Zillien N, Hargittai E (2009) Digital distinction: Status-specific types of Internet usage. *Soc. Sci. Quart.* 90(2):274–291.