Videogames and the Gender Gap in Computer Science

Claudia Meza-Cuadra*

Job Market Paper. Click here for the latest version

November 15, 2023

Abstract

In contrast to other STEM fields, over the last thirty years computer science has grown increasingly male-dominated. Using large-scale US survey data on field of education and occupation as well as data on computer and videogame playing from the mid-1980s to the early 2000s, I study the effect of the spread of videogames on the widening of the gender gap in computer science. Using two-way fixed effects regressions exploiting variation in the spread of videogames by birth state and cohort, I find that for men, greater exposure to videogames when young is associated with (1) an increased probability of obtaining a bachelor's degree in computer science and (2) an increased probability of working in a computer-related occupation. For those who obtain a bachelor's degree, a 20% increase in the spread of videogames in their cohort (while teenagers) is associated with a 10% increase in the gender gap in the probability of studying computer science, and a 14% increase in the gender gap in the probability of working in a related field. To address potential endogeneity, I instrument for videogame exposure using the early prevalence of videogame arcades by state combined with national sales of games released each year. The IV analysis confirms the results, suggesting that videogame exposure may be an important driver of the gender gap in computer science, and providing additional evidence of the long-term role of non-academic activities during childhood on education and career choices and outcomes.

*Universitat Pompeu Fabra, Department of Economics. E-mail: claudia.mezacuadra@upf.edu.

1 Introduction

Despite being approximately half of the US workforce, women held only 35% of jobs in STEM occupations in 2020 (National Science Foundation, 2020). This gender gap is even more pronounced in computer science, where women earned only 21% of bachelor's degrees, a smaller share than in 1995. Qualitative evidence suggests that young men's greater experience with computers before college may be one factor behind their greater participation in computer science education (Margolis and Fisher 2002). In the last 40 years, the popularity of video games has surged, particularly among young men, who spent on average almost 2 hours per day on gaming in 2022 (BLS, 2023). The ubiquity of video games means that they have potentially an unprecedented influence on the interests, skills, and aspirations of young people.

In this paper, I study the effect of the spread of videogames on the widening of the gender gap in computer science. To do this, I combine large-scale US survey data on field of education and occupation from the American Community Survey (ACS) with survey data on videogame playing from the mid-1980s to the early 2000s to measure the popularity of videogames in a person's place of birth. In the main analysis I estimate two-way fixed effects regressions using variation in the spread of videogames in the respondent's state of birth during their childhood. In line with my hypotheses, I show that for men, greater exposure to videogames when young is associated with (1) an increased probability of obtaining a bachelor's degree in computer science and (2) an increased probability of working in a computer-related occupation. For those who obtain a bachelor's degree, a 20% increase in the spread of videogames in their cohort (while teenagers) is associated with a 10% increase in the gender gap in the probability of studying computer science, and a 14% increase in the gender gap in the probability of working in a related field.

Exposure to videogames may be correlated with other characteristics that vary by state and time, such as local economic conditions, employment in and prevalence of technology and related industries, or educational quality, which could directly affect demand for computer science education and work. To address this potential endogeneity, I use a shift-share instrument for videogame exposure constructed using per capita density of videogame arcades by state combined with total North American sales of games released in each year. Arcades were early physical locations where young people played videogames, and spread quickly throughout the US in the 70s and 80s. I argue that the early distribution of arcades was not (only) driven by demand factors. For example, parents and legislators concerned about the negative effects of videogames on children as well as dangerous influences in public arcades attempted to and succeeded in passing bans or restrictions on arcade locations in several states including Georgia, Rhode Island, New Hampshire, New York, California, Texas and Massachusetts (Kocurek 2012). As such, the density of arcades is likely to affect the exposure of young people to videogames but to not be directly related to other local characteristics. Moreover, there is evidence that arcades were highly male dominated spaces (Williams 2006). My IV analysis confirms a significant effect for boys more exposed to videogames on the probability of obtaining a bachelor's degree in computer science, as well as working in a related field.

This paper contributes to the literature on the differences in how men and women choose their field of study and the causes of occupational segregation. Although women have closed the gap in human capital accumulation, large gender differences in occupation and industry remain (Cortes and Pan 2018, Blau and Kahn 2017). Researchers have pointed to several factors behind this persistent sorting, including discrimination, workplace flexibility, and environment, as well as differences in preferences and personality traits, gender identity and social norms (Bertrand and Duflo 2017, Goldin 2014, Bertrand 2011)

I provide additional evidence on the long-term role of non-academic activities during childhood and cultural influences on education and career choices. This is particularly relevant given the frequently gendered nature of entertainment and online media.

I also contribute to the extensive literature investigating the persistent gender gap in STEM in higher education and the labor force ((Kahn and Ginther 2017, Reuben et al. 2015, Ginther and Rosenbloom 2018). As noted by Kahn and Ginther (2017), women's underrepresentation in STEM bachelor's degrees continues to this day, and is concentrated in a subset of fields including physical and earth sciences, engineering, math, economics, and computer science. Of these fields, computer science has been a source of particular interest due to the sharp increase in the share of men obtaining bachelor's degrees in the field starting in the mid-1980's ((Kahn and Ginther 2017, Ginther and Rosenbloom 2018, Card and Payne 2021)). This paper contributes to understanding the factors behind this growing gender gap, providing quantitative evidence that videogames may have played a role.

Finally, this paper is related to the literature on the effects of video game use and its relationship to gender. Research has found that the use of videogames can improve spatial skills and problem-solving skills (Spence and Feng 2010, Barr 2017, Subrahmanyam and Greenfield 1994)).

On the other hand, there is some evidence that time spent playing games may crowd-out time spent on other learning activities, and may have negative effects on academic performance (Weis and Cerankosky 2010). Studies have also found that men are more likely to play videogames than women, and that those who play spend more time doing so ((Engelstätter and Ward 2022, Aguiar et al. 2021). Moreover, research has found strong gender biases in videogame content and advertising, with a greater proportion of characters being male, and female characters often depicted in sexualized and gender stereotyped ways (Scharrer 2004, Downs and Smith 2010, Bègue et al. 2017, Williams 2006). In addition, women who play online videogames often report experiencing sexual harassment (Kuznekoff and Rose 2013, Tang and Fox 2016). I build on this research by showing that videogame use may have lasting effects on academic and professional outcomes that differ by gender.

This paper proceeds as follows. In section 2 describe my data and the background for my instrumental variable. Then, in section 3 I present the empirical strategy in detail. Section 4 presents the main results, while section 5 discusses a series of robustness checks, and section 6 concludes.

2 Data

The main analysis is conducted using two main sources of data. First, I obtain information on education and occupation from the American Community Survey (ACS), for every year from 2009 to 2021. Data from the ACS is representative at the state level, and includes detailed questions about individuals' educational attainment, field of bachelor's degree, and field of occupation at the time of the survey. To isolate the effect on degree choice, I limit the sample to individuals aged 22 and over at the time of the survey, the minimum age at which individuals in the data report obtaining a degree. I obtain a sample of over 9.4 million individuals born between 1967 and 1994. Table 1a shows summary statistics for this sample.

In the main sample, only 36% of individuals had obtained at least a bachelor's degree. Moreover, around 1% of the sample had obtained a bachelor's degree in computer science. The average videogame exposure was 31% and approximately 2% of individuals listed their occupation as being computer related.

To further isolate the effect on the choice of college major, I also run all of the analyses including

Table 1a: Summary statistics for ACS full sample

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Survey year	9,403,062	2015	3.538	2009	2021
Birth year	9,403,062	1980	7.748	1967	1994
Male	9,403,062	0.501	0.500	0	1
Bachelor's	9,403,062	0.355	0.479	0	1
Computer science bachelor's	9,403,062	0.0103	0.101	0	1
Videogame exposure	9,403,062	0.306	0.207	0.00775	0.810
Computer-related occupation	8,488,112	0.0223	0.148	0	1

The full sample includes people born between 1967 and 1994 who were interviewed in the American Community Survey at ages 22 and above in 2009-2019 or 2021. Birth cohorts are defined in 4-year intervals between 1967 and 1994. Videogame exposure is the share of each cohort playing videogames at ages 10-17 in a person's birth state is calculated using CPS waves in 1984, 1989, 1993, 1997, 2001 and 2003.

only individuals who obtained (at least) a college degree. Summary statistics for this sample are shown in table 1b. In this smaller sample, 44% are men while computer science bachelor's are obtained by nearly 3% of individuals. Average videogame exposure is around 30%, as in the main sample, while almost 4% of people worked in computer-related occupations.

Table 1b: Summary statistics for ACS college sample

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Survey year	2,643,435	2014	3.520	2009	2021
Birth year	2,643,435	1979	7.416	1967	1994
Male	2,643,435	0.440	0.496	0	1
College	2,643,435	1	0	1	1
Computer science bachelor's	2,643,435	0.0289	0.168	0	1
Videogame exposure	2,643,435	0.297	0.203	0.00775	0.810
Computer-related occupation	2,643,435	0.0388	0.193	0	1

The college sample includes people born between 1967 and 1994 who were interviewed in the American Community Survey at ages 22 and above in 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017 or 2021 and had obtained a bachelor's degree at any point. Videogame exposure is the share of a person's cohort playing videogames at ages 10-17 in a person's birth state is calculated using CPS waves in 1984, 1989, 1993, 1997, 2001 and 2003.

Next, I construct a measure of exposure to videogames by state of residence and year of birth using micro-data from the CPS Computer and Internet Use Supplement. All individuals in households included in the survey are asked whether they use computers at home to play games at the time of the survey. To identify exposure to videogames before individuals apply to college, I limit the sample to individuals who were between 10 and 17 years old at the time of the survey. I then create the exposure variable as the fraction of the sample born in each year and state that reported playing videogames. Pooling data from 1984, 1989, 1993, 1997, 2001 and 2003, I obtain a sample of 101,439 individuals born between 1967 and 1993. Table 1c shows summary statistics for this sample. The sample is 51% male, and 34% report using their computer at home to play videogames.

Table 1c: Summary statistics for CPS sample

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Survey year	101,439	1994	6.794	1984	2003
Age	101,439	13.48	2.283	10	17
Year of birth	101,439	1981	7.193	1967	1993
Male	101,439	0.511	0.500	0	1
Plays videogames	101,439	0.345	0.475	0	1

The full sample includes people who were ages 10 to 17 and part of households interviewed in the CPS in 1984, 1989, 1993, 1997, 2001 and 2003. The variable "plays videogames" is an indicator variable equal to one if an individual answered yes when asked if they use computers at home to play games at the time of the survey.

To obtain a measure of videogame exposure when young for individuals in the ACS, I connect the two datasets using year and state of birth.* I also conduct the analyses at a higher level of geographic disaggregation, by connecting the two datasets using the metropolitan area of residence at the time of each survey. However, due to the high rates of mobility within the US, this may introduce bias into the results as people are likely to select into locations with greater labor market opportunities.

2.1 Identification: Amusement arcades

Exposure to videogames may be correlated with other characteristics that vary by state and time, such as local economic conditions, employment in and prevalence of technology and related industries, or educational quality, which could directly affect demand for computer science education and work. To address thispotential endogeneity, I use a shift-share instrument for videogame exposure constructed using per capita density of videogame arcades by state combined with total North American sales of games released in each year.

Arcades were early physical locations where young people played videogames, and spread quickly throughout the US in the 70s and 80s. As arcades became increasingly popular and

*The implied assumption is that individuals are likely to have spent their teenage years in their state of birth. In the ACS data, over 80% of teenagers lived in their state of birth.

common, they also became a source of controversy, and numerous communities organized against them, passing zoning ordinances and other restrictions (Kocurek 2012). These restrictions were motivated by concerns over the negative effects of videogames on children, as well as by a general association of the arcade business with gambling and crime (Kocurek 2012, Skolnik and Conway 2019).

During this time, video game arcades were highly male dominated spaces (Kaplan 1983, Skolnik and Conway 2019)). This may in part be due to the association between early video arcades and older coin-operated amusement machines which often included female nudity and sexual content, and were often placed in male-populated venues (e.g. bars and pool halls) (Kocurek 2012, Skolnik and Conway 2019). When women did visit arcades, they were more likely to play non-violent games, however these types of games were less common. Moreover, there is some evidence that parents were less likely to allow their daughters to frequent arcades (Kaplan 1983, ELLIS 1984).

In order to expand their client base and reduce the risk of further restriction, in the 1980s the amusement industry tried to recast arcades as wholesome entertainment centers for the entire family. New, often chain businesses, arcades were brightly-lit and located in shopping malls (Kaplan 1983, Skolnik and Conway 2019). Video game arcade machines could also be found in other businesses such as movie theaters, bowling alleys, pizza restaurants, grocery stores, pharmacies, etc. (Kocurek 2012). Despite continuing to be a minority at arcades, teenage girls and women did visit local arcades at a higher rate following the release of Pac-man, a game specifically designed to attract them (Newman 2018).

Given their popularity in this early period, the density of arcade machines likely affected the exposure of young people, especially boys, to videogames. Moreover, there is some evidence that children living in proximity to arcades are more likely to spend time there (Fisher 1995). Given their usual location in shopping malls and variation in their frequency due to local restrictions, their density is likely to be partially exogenous to education and labor market outcomes, after including controls.

In order to construct the instrument I first obtain a list of addresses of businesses which had arcade machines operating in the United States in the late 1980s from an industry trade magazine. I combine this with population in 1990 from the US Census Bureau to calculate the number of arcades by state per capita (per hundred thousand residents). Figure 2.1 shows a map and figure 2.1, a histogram of this variable. As shown in the figures, the number of arcades per 100,000 by state ranged from less than one to over 8.

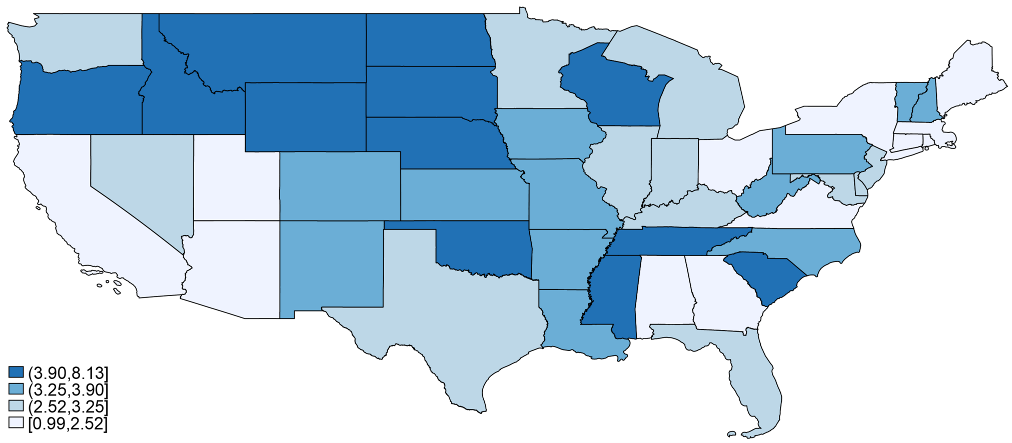

Figure 1: Map of arcades per capita in 1990

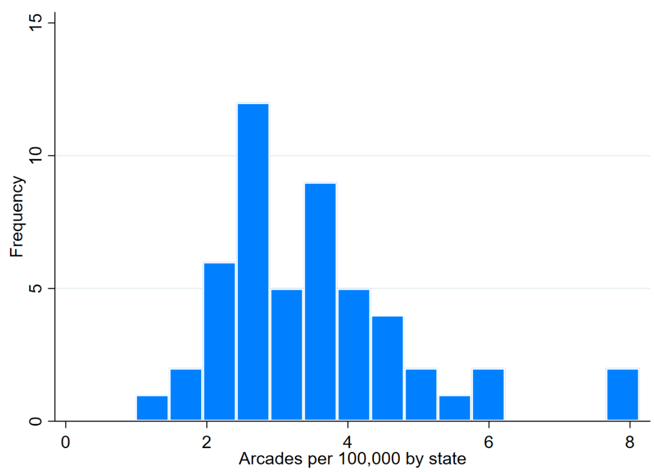

Figure 2: Frequency distribution of arcades per capita in late 1980s

I combine this measure of relative arcades per capita with a measure of the popularity of videogames released in each year. I aggregate sales by year of release from a database of sales to consumers of every videogame released in each year from VGChartz (`http://www.vgchartz.com/gamedb/`). This dataset was developed by VGChartz using data from a representative sample of small retailers combined with statistical prediction methods (VGChartz 2023) and has been used in recent research (Suziedelyte 2021). A limitation of this data is that yearly sales are not available for the study period, however the majority of sales usually occur within the first two years following the release of a game, meaning that all time-sales are likely to reflect the popularity of games released in any given year (Suziedelyte 2021).

3 Empirical Strategy

To investigate the role of videogames on the gender gap in computer science I start by using a two way fixed effects model which exploits variation in exposure to videogames by state and year of birth.

For each outcome I run the following regression at the individual level

$$
\begin{aligned}
Y_{is\tau t} = \alpha + \beta_1 videogames_{s\tau} + \beta_2 male_i + \beta_3 videogames_{s\tau} \times male_i + \\
X_{s\tau} + \mu_s \times \tau + \mu_s + \lambda_\tau + \delta_t + \epsilon_{is\tau t} \qquad (1)
\end{aligned}
$$

where for each individual i observed in survey year t, $videogames_{s\tau}$ is the share of people in their four-year birth cohort τ and birth state s observed at ages 10-17 who report using a computer at home to play videogames, β_3 is the coefficient of interest. The outcome Y_{ist} is an indicator for studying a computer-related field in college or working in a related occupation. I cluster standard errors in all regressions at the state-cohort level, as this is the level at which the main treatment variable is constructed.

In the main specification, I include as controls at the state cohort level the share of adults aged 22-65 who had a college degree, as well as mean and median household income. It is also important to include state fixed effects, μ_s, to control for time-invariant characteristics of a location which could impact both videogame exposure and a person's education and occupational outcomes. Similarly, I include cohort fixed effects, λ_τ, to capture the effect of changes in technology or other characteristics across time. Survey year fixed effects, δ_t, capture differences in educational and labor market decisions which may be due to the year in which the interview was conducted. In one specification I also include state specific time trends, $\mu_s \times \tau$, aggregating years at the cohort level, to capture differences in state characteristics over time unrelated to the spread of videogames.

The main analyses use variation in exposure to videogames at the state-cohort level. A greater level of geographic disaggregation is not available for individuals' past residence or location of birth. However, the results are robust to running the analyses at the metropolitan area level using individuals' location of residence at the time of the ACS survey. In this robustness analysis I also limit the sample to individuals who have always lived in the same state, to partially address potential selection into location of residence.

Exposure to videogames is measured at ages 10-17 in order to maximize the period of potential exposure prior to the age at which people usually apply to and begin higher education.

Moreover, the data on videogame use show that young people play the most games around the age of 15. The results are robust to using a shorter period of exposure from ages 12-15. Likewise, I run the main regressions using 4-year cohorts by year of birth, however the results are robust to using smaller cohorts. Figure 3 shows the average rate of videogame exposure for individuals in each cohort, where cohort 1 are individuals born between 1967 and 1970 and cohort 7, those born between 1991 and 1994.

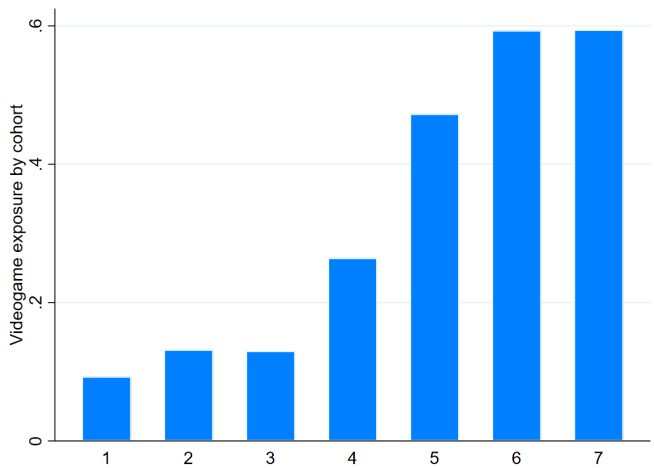

Figure 3: Average videogame exposure for individuals in each cohort

Next, to address potential endogeneity in the measure of videogame exposure I use an instrumental variable approach. The videogame exposure instrument is a shift-share which aims to capture the relative popularity of videogames in each state when individuals were teenagers. The instrument is constructed as follows:

$$exposureIV_{s\tau} = NAsales_{\tau} \times \frac{1}{\sum_{r \in S} \frac{arcades_{r,1990}}{population_{r,1990}}} \times \frac{arcades_{s,1990}}{population_{s,1990}} \tag{2}$$

where $NAsales_{\tau}$ is average all-time North American sales of videogames released in the years that individuals in cohort τ were aged 10-17 and $arcades_s$ is number of videogame arcades in state s in 1990 divided by population. This generates a variable that distributes annual sales of videogames across states according to the relative number of arcades that were open in that state in 1990, normalized by population.

The validity of my instrumental strategy relies on the idea that the per-capita density of arcades affects educational and occupational outcomes only through its effect on local exposure to videogames during a person's youth, after including controls and controlling for state, cohort and survey year fixed effects (Goldsmith-Pinkham et al. 2020).

First-stage. Using state and year of birth, I first assign videogame exposure and the exposure IV to each individual in the ACS sample. Given that the measure of videogame exposure is used both directly and interacted with an indicator for male, it is necessary to instrument for both variables. Thus I estimate the two following first-stage equations at the individual level:

$$\begin{aligned} videogames_{s\tau} =& \gamma_1 + \gamma_2 exposureIV_{s\tau} + \gamma_3 male_i + \beta_3 exposureIV_{s\tau} \times male_i + X_{s\tau} + \mu_s + \\ & \lambda_\tau + \delta_t + \epsilon_{ist\tau} \end{aligned} \tag{3}$$

$$\begin{aligned} videogames_{s\tau} \times male_i =& \gamma_1 + \gamma_2 exposureIV_{s\tau} + \gamma_3 male_i + \beta_3 exposureIV_{s\tau} \times male_i + X_{s\tau} + \mu_s + \\ & \lambda_\tau + \delta_t + \epsilon_{ist\tau} \end{aligned} \tag{4}$$

where $videogames_{s\tau}$ is the share of people in a person's birth cohort τ and birth state s observed at ages 10-17 who played videogames, $male_i$ is an indicator for men, $exposureIV_{s\tau}$ is the arcade and sales based exposure instrument, $X_{s\tau}$ are state-cohort level controls, μ_s is state fixed effects, λ_τ is cohort fixed effects, and δ_t is survey year fixed effects. Standard errors are clustered at the state by cohort level.

My instrumental strategy is thus to use both $exposureIV_{s\tau}$ and $exposureIV_{s\tau}$ interacted with a dummy for male as instruments. In order to be valid, this requires gender to be exogenous, a plausible assumption which I verify.

4 Results

Only 1% of individuals in the full sample obtained bachelor's degrees in computer science, however 78% of those who did were men. Table 2 shows that even after including controls, there was a 1pp gender gap in the probability of obtaining a degree in CS. Focusing on the sample of people who obtained a bachelor's degree, this gap is larger, with men being 3.5pp more likely to study computer science.

Table 2: Probability of studying a computer-related field, for full sample and college sample

	(1)	(2)	(3)	(4)
VARIABLES	Full sample		College sample	
Male	0.0104***	0.0104***	0.0351***	0.0351***
	(0.000383)	(0.000383)	(0.00105)	(0.00105)
Videogame exposure	-0.00127	-0.00643***	-0.00982**	-0.0164***
	(0.00185)	(0.00233)	(0.00467)	(0.00601)
Male × Videogame exposure	0.00417***	0.00417***	0.0189***	0.0188***
	(0.000992)	(0.000992)	(0.00244)	(0.00244)
Observations	9,403,062	9,403,062	3,315,886	3,315,886
Cohort FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes
State time trend	No	Yes	No	Yes

Note. The outcome is a variable equal to 1 if a person obtained a bachelor's degree in a computer-related field and 0 otherwise. Videogame share is the share of a person's four-year birth cohort that was playing videogames at ages 10-17 in a person's birth state calculated using CPS waves in 1984, 1989, 1993, 1997, 2001 and 2003. Standard errors clustered at the state x cohort level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

In line with my hypothesis, more exposure to videogames is related to a larger gender gap in the probability of obtaining a bachelor's degree in computer science. Specifically, the OLS results in Table 2 show that boys with more videogame exposure during adolescence are more likely to study a computer related field than similar girls. In contrast, if anything, the effect of the spread of videogame exposure on girls is negative (coefficient on videogame exposure). The effects observed in the full sample are modest but significant and robust to including fixed effects and state specific time-trends.

The average person in the full sample was born in a state and cohort in which 30% of teenagers played videogames. A one standard deviation or 20 percentage point increase in the rate of videogame exposure is associated with a 7% larger gender gap in computer science at the undergraduate level or an 11% larger gap when limiting the sample to those who obtained a bachelor's degree.

I obtain similar results when instrumenting for videogame exposure, as shown in Table 3. The coefficients on videogame exposure are in the same direction as in the OLS regresssions, with a negative effect on exposure but a positive effect on the interaction with the dummy for male. In the full sample, the estimated magnitude of the effect on the gender gap is larger, with a one standard deviation increase in the rate of videogame exposure associated with a 13% increase in the gap. Limiting the sample to those who obtained a bachelor's degree increases the estimated effect to 17%.

Table 3: Probability of studying a computer-related field, instrumenting for videogame exposure

	(1) Full sample	(2) Full sample	(3) College sample	(4) College sample
Male	0.00943***	0.00942***	0.0325***	0.0324***
	(0.000384)	(0.000388)	(0.000917)	(0.000920)
Videogame exposure	-0.0579	-0.106	-0.129**	-0.191
	(0.0362)	(0.0884)	(0.0587)	(0.118)
Male × videogame exposure	0.00743***	0.00746***	0.0275***	0.0276***
	(0.00137)	(0.00138)	(0.00302)	(0.00302)
Observations	9,403,062	9,403,062	3,315,886	3,315,886
Cohort FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes

Note. The videogame exposure instrument is equal to average north American sales of videogames released when the members of each cohort were ages 10-17, distributed by state according to each state's total number of arcades in 1990, multiplied by 100,000 and divided by population. The outcome is a variable equal to 1 if a person obtained a bachelor's degree in a computer-related field and 0 otherwise. Videogame share is the share of a person's four-year birth cohort that was playing videogames at ages 10-17 in a person's birth state calculated using CPS waves in 1984, 1989, 1993, 1997, 2001 and 2003. Standard errors clustered at the state x cohort level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4 shows the results on field of occupation. In the full sample, 2% of individuals were employed in a computer-related occupation. As shown in Table 4, after including controls, men were 2 percentage points more likely than women to work in this field, with a larger gender gap among those who obtained bachelor's degrees. As with education, exposure to videogames is associated with the likelihood that an individual works in a computer-related occupation. Morevoer, the coefficients on videogame exposure are negative, while those on its interaction with the dummy for male are consistently positive indicating an effect on increasing the gender gap. However, the estimated effects on the full sample are smaller. Specififcally, a one standard deviation or 20 percentage point increase in the rate of videogame exposure is associated with only a 3% increase in the size of the gender gap. If people who did not obtain bachelor's degrees are excluded, this increases to 13%.

Table 4: Probability of working in a computer-related occupation, for full sample and college sample

VARIABLES	(1) Full sample	(2)	(3) College sample	(4)
Male	0.0208***	0.0208***	0.0397***	0.0397***
	(0.000589)	(0.000589)	(0.000858)	(0.000859)
Videogame exposure	-0.00185	-0.00612***	-0.0232***	-0.0213***
	(0.00201)	(0.00209)	(0.00408)	(0.00415)
Male × videogame exposure	0.00548***	0.00548***	0.0292***	0.0291***
	(0.00145)	(0.00145)	(0.00224)	(0.00224)
Observations	8,488,112	8,488,112	3,190,760	3,190,760
Cohort FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes
State time trend	No	Yes	No	Yes

Note. The outcome is a variable equal to 1 if a person lists their main occupation as being computer-related. The sample is limited to people who work or have worked in the last 5 years. Videogame share is the share of a person's four-year birth cohort that was playing videogames at ages 10-17 in a person's birth state calculated using CPS waves in 1984, 1989, 1993, 1997, 2001 and 2003. Standard errors clustered at the state x cohort level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Once again, the 2sls estimates suggest a larger effect of videogame exposure on the gender gap than the OLS estimates (Table 5). While the effect of videogame exposure itself becomes insignificant once controls are included, the estimated coefficient on the interaction with the male dummy is positive and strongly significant. In this case, a 20 percentage point increase in the rate of videogame exposure is associated with 9% increase in the gap in the probability that men are occupied in a computer-related field, compared to women. For those in the college sub-sample, this increase in videogame exposure is associated with a 19% larger gap.

Overall, given the increase in videogame exposure from 13% to 59% between the oldest and youngest cohorts, the results suggest that videogames could explain a significant part of the growth in the gender gap in computer science education throughout this period.

In order to better understand differences in the magnitude of estimated effects in the two samples, I re-run the analyses using an indicator for obtaining a bachelor's degree as the outcome. Table 6 shows the results. After including controls, men are over 6 percentage points less likely to obtain a bachelor's degree than women in the sample. Moreover, the results suggest that the effect of videogame exposure on this probability is in the opposite direction to the effects on computer science. In particular, videogame exposure is positively associated with the probability of attending college for women, but negatively associated with with this outcome for men. The

Table 5: Probability of working in a computer-related field, instrumenting for videogame exposure

	(1)	(2)	(3)	(4)
	Full sample	Full sample	College sample	College sample
Male	0.0192***	0.0192***	0.0374***	0.0374***
	(0.000583)	(0.000586)	(0.000822)	(0.000826)
Videogame exposure	-0.0652*	-0.110	-0.216**	-0.285
	(0.0391)	(0.0851)	(0.0919)	(0.175)
Male × videogame exposure	0.0107***	0.0107***	0.0364***	0.0365***
	(0.00178)	(0.00179)	(0.00268)	(0.00269)
N	8488112	8488112	3190760	3190760
Cohort FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes

Note. The sales instrument is equal to average north American sales of videogames released when the members of each cohort were ages 10-17, distributed by state according to each state's total number of arcades in 1990, multiplied by 100,000 and divided by population. The outcome is a variable equal to 1 if a person lists their main occupation as being computer-related. The sample is limited to people who work or have worked in the last 5 years. Videogame share is the share of a person's four-year birth cohort that was playing videogames at ages 10-17 in a person's birth state calculated using CPS waves in 1984, 1989, 1993, 1997, 2001 and 2003. Standard errors clustered at the state x cohort level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

estimated effects suggest that a one standard deviation increase in exposure to videogames would result in an increase in the gender gap from 6% to 7.4%.

Table 6: Probability of getting a bachelor's degree

	(1)	(2)
VARIABLES		Bachelor's
Male	-0.0627***	-0.0627***
	(0.00270)	(0.00270)
Videogame exposure	0.110***	0.0243*
	(0.0222)	(0.0124)
Male × videogame exposure	-0.0733***	-0.0734***
	(0.00574)	(0.00574)
Observations	9,338,689	9,338,689
Cohort FE	Yes	Yes
State FE	Yes	Yes
State time trend	No	Yes

Note. The outcome is a variable equal to 1 if a person obtained a bachelor's degree and 0 otherwise. Videogame share is the share of a person's four-year birth cohort that was playing videogames at ages 12-15 in a person's birth state calculated using CPS waves in 1984, 1989, 1993, 1997, 2001 and 2003. Standard errors clustered at the state x cohort level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Next, I re-do this analysis instrumenting for videogame exposure, shown in table 7. Reassuringly, the coefficient on videogame exposure interacted with the dummy for male is very similar to that in the OLS regressions. On the otherhand, the average effect of videogame exposure becomes insignificant, suggesting that there may be endogeneity in this variable which is addressed through the instrumental variable.

Table 7: Probability of getting a bachelor's degree, instrumenting for videogame exposure

	(1)	(2)
	Full sample	Full sample
Male	-0.0620***	-0.0621***
	(0.00234)	(0.00235)
Videogame exposure	-0.271	-0.895
	(0.357)	(0.860)
Male × videogame exposure	-0.0758***	-0.0754***
	(0.00594)	(0.00599)
Observations	9338689	9338689
Cohort FE	Yes	Yes
State FE	Yes	Yes
Controls FE	No	Yes

Note. The videogame exposure instrument is equal to average north American sales of videogames released when the members of each cohort were ages 10-17, distributed by state according to each state's total number of arcades in 1990, multiplied by 100,000 and divided by population. The outcome is a variable equal to 1 if a person obtained a bachelor's degree and 0 otherwise. Videogame share is the share of a person's four-year birth cohort that was playing videogames at ages 10-17 in a person's birth state calculated using CPS waves in 1984, 1989, 1993, 1997, 2001 and 2003. Standard errors clustered at the state x cohort level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

These results suggest that videogame exposure may have two opposing effects on men: first, men more exposed to games are less likely to obtain a bachelor's degree, perhaps shifting into more technical occupations, and second, those who do obtain university degrees are more likely to study computer science.

5 Robustness

I conduct a wide array of robustness checks. First, I re-do the analysis using videogame exposure measured during a shorter age-range, using individuals observed at ages 12-15 instead of 10-17. The results shown in tables 8 and 9 are consistent with those of the main analysis, though the magnitude of the effects are smaller, likely due to the shorter period of exposure. Re-doing the analyses at the metropolitan area level of residence also gives similar results.

I also use alternative definitions of the instrumental variable. First, I conduct the analysis replacing the average North American sales per game with total sales. Next I construct the shift share instrument including only a subset of genres of videogames which have been identified to be particularly appealing to boys. In particular, I use sales of fighting, shooting and action games only. The results show a large effect of videogame exposure on the gender gap in the college sample.

Finally, one concern with the instrumental variable is that arcades are measured in the late 1980s, rather than earlier during the arcade boom and may reflect demand for videogames as well. In order to address this concern, I obtain alternative data on the number of coin-operated

Table 8: Probability of studying a computer-related field, using exposure from ages 12-15

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES		Full sample			College sample	
Male	0.0109***	0.0109***	0.0109***	0.0365***	0.0367***	0.0367***
	(0.000452)	(0.000452)	(0.000451)	(0.00131)	(0.00130)	(0.00130)
Videogame exposure	-0.00447***	0.000845	0.000375	-0.0115***	-0.00259	-0.00187
	(0.000421)	(0.00171)	(0.00175)	(0.000913)	(0.00457)	(0.00457)
Male × videogame exposure	0.00203*	0.00210*	0.00210*	0.0138***	0.0133***	0.0132***
	(0.00113)	(0.00114)	(0.00114)	(0.00297)	(0.00296)	(0.00295)
Observations	6,786,694	6,786,694	6,786,694	2,369,862	2,369,862	2,369,862
Cohort FE	No	Yes	Yes	No	Yes	Yes
State FE	No	Yes	Yes	No	Yes	Yes
State time trend	No	No	Yes	No	No	Yes

Note. The outcome is a variable equal to 1 if a person obtained a bachelor's degree in a computer-related field and 0 otherwise. Videogame share is the share of a person's four-year birth cohort that was playing videogames at ages 12-15 in a person's birth state calculated using CPS waves in 1984, 1989, 1993, 1997, 2001 and 2003. Standard errors clustered at the state x cohort level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Probability of working in a computer-related occupation, using exposure from ages 12-15

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES		Full sample			College sample	
Male	0.0220***	0.0220***	0.0220***	0.0418***	0.0417***	0.0418***
	(0.000658)	(0.000661)	(0.000661)	(0.000998)	(0.00100)	(0.00100)
Videogame exposure	-0.00405***	0.000512	-0.000612	-0.00270**	-0.0138***	-0.0104***
	(0.000641)	(0.00171)	(0.00170)	(0.00114)	(0.00388)	(0.00323)
Male × Videogame exposure	0.00168	0.00166	0.00166	0.0218***	0.0217***	0.0216***
	(0.00153)	(0.00153)	(0.00153)	(0.00248)	(0.00249)	(0.00249)
Observations	6,142,797	6,142,797	6,142,797	2,284,017	2,284,017	2,284,017
Cohort FE	No	Yes	Yes	No	Yes	Yes
State FE	No	Yes	Yes	No	Yes	Yes
State time trend	No	No	Yes	No	No	Yes

Note. The outcome is a variable equal to 1 if a person lists their main occupation as being computer-related. The sample is limited to people who work or have worked in the last 5 years. Videogame share is the share of a person's four-year birth cohort that was playing videogames at ages 12-15 in a person's birth state calculated using CPS waves in 1984, 1989, 1993, 1997, 2001 and 2003. Standard errors clustered at the state x cohort level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 10: Probability of studying a computer-related field, using 2-year cohorts

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES		Full sample			College sample	
Male	0.0109***	0.0109***	0.0109***	0.0366***	0.0368***	0.0368***
	(0.000418)	(0.000418)	(0.000418)	(0.00118)	(0.00117)	(0.00117)
Videogame exposure	-0.00468***	-0.00659***	-0.00904***	-0.0121***	-0.0179***	-0.0224***
	(0.000405)	(0.00157)	(0.00157)	(0.000875)	(0.00385)	(0.00387)
Male × videogame exposure	0.00229**	0.00236**	0.00236**	0.0143***	0.0137***	0.0137***
	(0.00109)	(0.00110)	(0.00109)	(0.00271)	(0.00270)	(0.00270)
Observations	6,623,530	6,623,530	6,623,530	2,322,050	2,322,050	2,322,050
Cohort FE	No	Yes	Yes	No	Yes	Yes
State FE	No	Yes	Yes	No	Yes	Yes

Note. The videogame exposure instrument is equal to average north American sales of videogames released when the members of each cohort were ages 10-17, distributed by state according to each state's total number of arcades in 1990, multiplied by 100,000 and divided by population. The outcome is a variable equal to 1 if a person obtained a bachelor's degree in a computer-related field and 0 otherwise. Videogame share is the share of a person's 2-year birth cohort that was playing videogames at ages 10-17 in a person's birth state calculated using CPS waves in 1984, 1989, 1993, 1997, 2001 and 2003. Standard errors clustered at the state x cohort level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Probability of working in a computer-related field, using 2-year cohorts

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES		Full sample			College sample	
Male	0.0220***	0.0219***	0.0219***	0.0220***	0.0219***	0.0219***
	(0.000492)	(0.000494)	(0.000494)	(0.000492)	(0.000494)	(0.000494)
Videogame exposure	-0.00405***	-4.49e-05	-0.000807	-0.00405***	-4.49e-05	-0.000807
	(0.000514)	(0.00135)	(0.00134)	(0.000514)	(0.00135)	(0.00134)
Male × Videogame exposure	0.00216*	0.00214*	0.00214*	0.00216*	0.00214*	0.00214*
	(0.00124)	(0.00124)	(0.00124)	(0.00124)	(0.00124)	(0.00124)
Observations	5,995,305	5,995,305	5,995,305	5,995,305	5,995,305	5,995,305
Cohort FE	No	Yes	Yes	No	Yes	Yes
State FE	No	Yes	Yes	No	Yes	Yes
State time trend	No	No	Yes	No	No	yes

Note. The outcome is a variable equal to 1 if a person lists their main occupation as being computer-related. The sample is limited to people who work or have worked in the last 5 years. Videogame share is the share of a person's two-year birth cohort that was playing videogames at ages 10-17 in a person's birth state calculated using CPS waves in 1984, 1989, 1993, 1997, 2001 and 2003. Standard errors clustered at the state x cohort level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

amusement businesses from the Quarterly Census of Employment and Wages (QCEW) (Bureau of Labor Statistics). Unlike the arcade location dataset, this includes only businesses which are dedicated to arcades. However, I verify that the number of arcades businesses in 1980 is strongly correlated with the original dataset and re-do the analysis using this data, obtaining consistent results.

Table 10: Probability of studying a computer-related field, using arcades in 1980

	(1)	(2)	(3)	(4)
	Full sample	Full sample	College sample	College sample
Male	0.0121***	0.0121***	0.0368***	0.0368***
	(0.000495)	(0.000505)	(0.00123)	(0.00122)
Videogame exposure	0.00247	0.112	-0.0108	-0.0710
	(0.0713)	(0.600)	(0.108)	(0.583)
Male × videogame exposure	-0.00126	-0.00130	0.0134***	0.0134***
	(0.00124)	(0.00125)	(0.00274)	(0.00269)
Observations	9403062	9403062	3315886	3315886
State & Cohort FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6 Conclusion

In this paper, I investigate the role of videogame in the widening of the gender gap in computer science. The results suggest that since the early 1980s, greater exposure to videogames has been associated with attracting a larger share of boys to study and work in computer-related fields. In particular, boys more exposed to games are significantly more likely than similar girls to have studied computer science. This suggests that the spread of videogames may have contributed to the growing and persistent gender gap in this field. Nevertheless, I find that videogames may have had complex effects on boys' educational outcomes, with those more exposed being less likely to obtain bachelor level degrees.

These results highlight the important role that non-academic activities can play in shaping young people's choices and economic outcomes. This is particularly relevant given the increasing popularity of online gaming, and ongoing debates about smartphone use in children and teenagers.

Moreover, the findings point to long-term impacts of differences in leisure-time and hobbies by gender during adolescence, which should be further explored. For instance, policymakers and parents concerned about gender segregation by occupation or to differences in economic outcomes by gender, should perhaps pay greater attention to children's non-academic interests and time-use.

References

Aguiar, M., Bils, M., Charles, K. K., and Hurst, E. (2021). Leisure Luxuries and the Labor Supply of Young Men. *Journal of Political Economy*, 129(2):337–382. Publisher: The University of Chicago Press.

Barr, M. (2017). Video games can develop graduate skills in higher education students: A randomised trial. *Computers & Education*, 113:86–97.

Bertrand, M. (2011). New Perspectives on Gender. In *Handbook of Labor Economics*, volume 4B, pages 1543–1590. Elsevier.

Bertrand, M. and Duflo, E. (2017). Chapter 8 - Field Experiments on Discrimination. In Banerjee, A. V. and Duflo, E., editors, *Handbook of Economic Field Experiments*, volume 1 of *Handbook of Field Experiments*, pages 309–393. North-Holland.

Blau, F. D. and Kahn, L. M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55(3):789–865.

Bègue, L., Sarda, E., Gentile, D. A., Bry, C., and Roché, S. (2017). Video Games Exposure and Sexism in a Representative Sample of Adolescents. *Frontiers in Psychology*, 8.

Card, D. and Payne, A. A. (2021). High School Choices and the Gender Gap in Stem. *Economic Inquiry*, 59(1):9–28. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecin.12934.

Cortes, P. and Pan, J. (2018). Occupation and Gender. ISBN: 9780190628963 Library Catalog: www.oxfordhandbooks.com.

Downs, E. and Smith, S. L. (2010). Keeping Abreast of Hypersexuality: A Video Game Character Content Analysis. *Sex Roles*, 62(11):721–733.

ELLIS, D. (1984). Video Arcades, Youth, and Trouble. *Youth & Society*, 16(1):47–65. Publisher: SAGE Publications Inc.

Engelstätter, B. and Ward, M. R. (2022). Video games become more mainstream. *Entertainment Computing*, 42:100494.

Fisher, S. (1995). The amusement arcade as a social space for adolescents: an empirical study. *Journal of Adolescence*, 18(1):71–86.

Ginther, D. K. and Rosenbloom, J. L. (2018). Why Do Women Leave Computer Science and Information Technology Jobs? In *Presentation at the Australian gender economics workshop*.

Goldin, C. (2014). A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, 104(4):1091–1119.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624. Publisher: American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Kahn, S. and Ginther, D. (2017). Women and STEM.

Kaplan, S. J. (1983). The Image of Amusement Arcades and Differences in Male and Female Video Game Playing. *The Journal of Popular Culture*, XVII(1):93–98. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.0022-3840.1983.1701_93.x.

Kocurek, C. A. (2012). Masculinity at the video game arcade : 1972-1983. Accepted: 2013-11-12T21:07:29Z.

Kuznekoff, J. H. and Rose, L. M. (2013). Communication in multiplayer gaming: Examining player responses to gender cues. *New Media & Society*, 15(4):541–556. Publisher: SAGE Publications.

Margolis, J. and Fisher, A. (2002). *Unlocking the clubhouse: Women in computing*. MIT press.

Newman, M. Z. (2018). *Atari Age: The Emergence of Video Games in America*. The MIT Press.

Reuben, E., Sapienza, P., and Zingales, L. (2015). Taste for Competition and the Gender Gap Among Young Business Professionals. Working Paper 21695, National Bureau of Economic Research. Series: Working Paper Series.

Scharrer, E. (2004). Virtual Violence: Gender and Aggression in Video Game Advertisements. *Mass Communication & Society*, 7(4):393–412. Place: US Publisher: Lawrence Erlbaum.

Skolnik, M. R. and Conway, S. (2019). Tusslers, Beatdowns, and Brothers: A Sociohistorical Overview of Video Game Arcades and the Street Fighter Community. *Games and Culture*, 14(7-8):742–762. Publisher: SAGE Publications.

Spence, I. and Feng, J. (2010). Video Games and Spatial Cognition. *Review of General Psychology*, 14(2):92–104. Publisher: SAGE Publications Inc.

Subrahmanyam, K. and Greenfield, P. M. (1994). Effect of video game practice on spatial skills in girls and boys. *Journal of Applied Developmental Psychology*, 15(1):13–32. Place: Netherlands Publisher: Elsevier Science.

Suziedelyte, A. (2021). Is it only a game? Video games and violence. *Journal of Economic Behavior & Organization*, 188:105–125.

Tang, W. Y. and Fox, J. (2016). Men's harassment behavior in online video games: Personality traits and game factors. *Aggressive Behavior*, 42(6):513–521. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ab.21646.

VGChartz (2023). Game database. Technical report.

Weis, R. and Cerankosky, B. C. (2010). Effects of Video-Game Ownership on Young Boys' Academic and Behavioral Functioning: A Randomized, Controlled Study. *Psychological Science*, 21(4):463–470. Publisher: SAGE Publications Inc.

Williams, D. (2006). Why Game Studies Now? Gamers Don't Bowl Alone. *Games and Culture*, 1(1):13–16. Publisher: SAGE Publications.

Appendix

Table A1 shows the two first stage regressions using the full sample. The outcome in column (1) is exposure to videogames while the outcome in column (2) is this variable interacted with a male dummy. The table shows that the interacted instrument is a particularly strong instrument. The F-statistic is large for column (2) but small for column (1). However, the Cragg-Donaldson statistic is 19,775 suggesting that the combined instruments are not weak.

Table A1: First stages

	(1) Average 1990: Videogame exposure × male	(2) Average 1990: Videogame exposure
Male	0.547***	0.000126**
	(0.0174)	(0.0000562)
ExposureIV	0.884***	-0.0826
	(0.0594)	(0.0730)
Male × ExposureIV	-1.868***	-0.00129***
	(0.110)	(0.000446)
N	9403062	9403062
F-statistic	213	3
State & Cohort FE	Yes	Yes
Controls	Yes	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2: First stages

	(1) Average 1980: Videogame exposure x male	(2) Average 1980: Videogame exposure
Male	0.452***	0.0000803*
	(0.0258)	(0.0000446)
ExposureIV	0.568***	0.0182
	(0.0788)	(0.0306)
Male × exposureIV	-1.145***	-0.000974***
	(0.167)	(0.000359)
N	9403062	9403062
State & Cohort FE	Yes	Yes
Controls	Yes	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$