APPENDIX TO: “Variety-Based Congestion in Online Markets: Evidence from Mobile Apps” (for Online Publication)

Daniel Ershov
UCL School of Management
d.ershov@ucl.ac.uk

A Appendix A

A.1 List of Non-Game Categories

Table A1: Google Play Non-Game Categories

Books & Reference	Libraries & Demo	Productivity
Business	Lifestyle	Shopping
Comics	Media & Video	Social
Communications	Medical	Sports
Education	Music & Audio	Tools
Entertainment	News & Magazines	Transportation
Finance	Personalization	Travel & Local
Health & Fitness	Photography	Weather

A.2 Consumer App Discovery Surveys

Two consumer surveys, taken by Forrester in 2012 (TechCrunch.com) and Google / Ipsos in 2014 (thinkwithgoogle.com 2014), asked app consumers how they discover new products. 58% of Android consumers discover new products through "General browsing in the app store", according to Forrester, and 25% discover new products through more targeted browsing - looking at "top rated" or "most popular" app lists in the app store (TechCrunch.com). Only a small share of consumers discover new apps through an internet search engine. Answers are similar in the Ipsos survey, with 40% of consumers browsing the app store to discover new products, and only 27% using search engines (thinkwithgoogle.com 2014).

Results from two additional app consumer surveys outside my sample period confirm these findings. In a Nielsen survey from 2011, over 60% of surveyed consumers on both Android and the iOS App Store stated they discover new products by searching the app store (BusinessInsider.com). This is not defined as "browsing" and could include using the search function of the store, but it is a far more popular method of discovery than advertising or other 3rd party sites. In another Google / Ipsos survey from 2016, once again, besides learning about apps from friends and family, the most popular discovery method is browsing the app store (thinkwithgoogle.com 2016). This last survey is from nearly two years after the end of my sample period, and it follows substantial improvements in the integration of apps into Google search results. Still, only 21% of users discover new apps using search engines.

B Appendix B

B.1 Data Management - Classifying App Types Using Text

I use a Random Forest machine learning algorithm that first maps the descriptions of the classified post-March 2014 apps into categories, and then projects this mapping on other apps. After removing "stopwords" (e.g., "and", "or") I convert app descriptions into vectors of words and then into term frequency-inverse document frequencies. This method assigns the highest weight to words appearing frequently in an app's description relative to the average description. I use April 2014 apps as the training set for a Random Forest classifier (other classifiers such as KNN give similar result). I then apply the classifier to apps in every month prior to March 2014. This is similar to how Liu et al. 2014 map Google Play categories into Apple iTunes categories.

B.2 Data Management - Predicting App Downloads

Raw app data includes a range of cumulative downloads that an app accrues over its lifetime. The full list of download ranges is in Table B1. This range is observable in every snapshot of the store. It is conceptually straight-forward to define "per-period downloads" as the difference in lifetime downloads between period t and period $t-1$. For example, the difference in the lower bounds of lifetime downloads, or in the average of lifetime downloads.

However, the bandwidth increases with the number of downloads, starting at 4 downloads ([1-5], [5-10]) and increasing to 40 ([10-50]) and eventually to 400 million ([100 million - 500 million]). This introduces two possible sources of measurement error, which become worse for more successful apps: (1) overestimation of per period downloads for apps that move from one level to another. An app with a range of [100 thousand - 500 thousand] downloads that moves to the [500 thousand- 1 million] range in the next period could have been downloaded 500 thousand times or 3 times. (2) underestimation of per period downloads for successful apps. An app in the [100 million - 500 million] download range can have millions of downloads every week and stay in the same range.

I rely on two features of the data to recover weekly or monthly app downloads. First, the lifetime download bandwidth for *new entrants* is equal to the per-period bandwidth: an app that entered one period ago and is in the 10 thousand to 50 thousand range was downloaded between 10 thousand and 50 thousand times in the period. Second, I observe weekly category rankings which reflect the 500 most-

Table B1: List of Cumulative Download Ranges

Lower Bound	Upper Bound
1	5
5	10
10	50
50	100
100	500
500	1,000
1,000	5,000
5,000	10,000
10,000	50,000
50,000	100,000
100,000	500,000
500,000	1 million
1 million	5 million
5 million	10 million
10 million	50 million
50 million	100 million
100 million	500 million
500 million	1 billion

downloaded apps in each category roughly over the past week.[1] At a weekly frequency, the rankings and downloads of new apps are known. Summary statistics are in Table B2.[2] I use these apps to predict the downloads of other apps in the market.

Several studies of online markets with best-seller lists find that the Pareto distribution accurately characterizes the rank-downloads relationship (Chevalier and Mayzlin 2006; Garg and Telang 2013).[3] The Pareto distribution is a negative ex-

[1]It is not precisely known how the lists are determined, but Google releases (AdWeek.com) as well as anecdotal industry evidence (Quora) suggest that they reflect the downloads of apps over the previous several days.

[2]I can assign the lower bound of the bandwidth as the number of weekly downloads, the upper part of the bandwidth, or the average of the bandwidth. In the rest of the analysis of this paper I assign the lower end of the bandwidth, since the average and upper parts of the bandwidth produce unrealistic estimates of downloads.

[3]It is possible that the Pareto distribution does not correctly predict downloads in this market. Eeckhout (2004) shows that the Pareto distribution accurately predicts the rank-size relationship for the upper tail of the distribution but not necessarily for the lower tail. Liebowitz and Zentner (2020) similarly shows evidence of potential inaccuracy in approximations using distributional assumptions.

ponential distribution where an app at rank n has exponentially more downloads than the app at rank $n+1$. I fit this distribution for every week and category by estimating an OLS regression of the logarithm of the rank of paid or non-paid ($p \in \{Paid, Non-Paid\}$) new app j in category c at week a of month t on the logarithm of the downloads for every category and week:

$$ln(\text{Downloads}_{jcat}) = \delta_{cp} + \delta_{tp} + \beta_1 ln(\text{Rank}_{jcat}) + \beta_2 ln(\text{Rank}_{jcat}) \times Paid_j + \mu_{jcat}$$

where δs are category and year/month dummies,[4] and where μ_{jcat} is a mean zero random variable representing measurement error. βs are slope coefficients that differ for paid and non-paid apps.[5] I use the lower bound of the bandwidth (minimum downloads in a week) as the dependent variable.[6] Summary statistics for new apps are in Table B2 and regression estimates are in Table B3. Estimated Pareto Distribution parameters are broadly consistent with similar exercises in the literature (Garg and Telang 2013; Leyden 2018).

I predict the downloads of all apps in the market with estimates from this regression. Only the top 500 ranks each week are observed. To generate rankings for the unranked apps, I sort them based on their number of cumulative lifetime downloads and their age in every week and break up ties by randomizing.[7]

This prediction algorithm depends on variation in app rankings over time. New apps should be able to enter into the rankings at different points in the distribution for me to estimate the Pareto relationship accurately. This is true in the data. While a large proportion of apps not change their rankings from week to week, many apps move at least two spots on a weekly basis. Figure B1 shows the distribution in weekly changes in game rankings.

I use an alternative measure of downloads that does not rely on the Pareto distribution in Online Appendix C.4.3. I also estimate my main results only for new apps, which are not affected by distributional assumptions. Results are qualitatively similar across the two sale proxies.

[4] δ_{cp} represents $\delta_{cp}\sum_c(D_c \times Paid_j + D_c \times (1 - Paid_j))$, where D_c is a dummy for whether j belongs to category c and $Paid_j$ is a dummy for whether app j is paid. δ_{tp} represents $\delta_{tp}\sum_t(D_t \times Paid_j + D_t \times (1 - Paid_j))$, where D_t is a dummy for month t.

[5] Predictions do not change substantially when slope coefficients also vary by time or category. I also test for the heterogeneity of the slope coefficients by app-ranking in Column (2) of Table B3, and do not find statistically significant differences in slope between very high ranked and lower ranked apps.

[6] Results using the upper bound or the average clearly overstate the number of downloads. For example, the model predicts each of the top 50 apps to have over 10 million weekly downloads.

[7] To check that randomization does not drive any of the main estimates, I re-estimate the analysis several times with different randomized seeds. The results remain qualitatively and quantitatively similar.

Table B2: **Summary Statistics of New Apps at Weekly Level**

Variable	Mean	Std. Dev.	Min	Max	N Obs
Games					
Download Lower Bound	19,035	202,837	1	1 million	15,958
Non-Games					
Download Lower Bound	8,473	322,478	1	5 million	28,699

Table B3: **Regression Results on Downloads**

Outcome Variable:	*ln(Min Downloads Bound)*		
	(1)	(2)	(3)
	Games		Non-Games
ln(Rank)	-0.973	-0.979	-0.981
	(0.074)	(0.080)	(0.046)
ln(Rank)×Paid	-1.170	-1.112	-1.016
	(0.060)	(0.070)	(0.084)
ln(Rank) × Low-Ranked		-0.010	
		(0.031)	
ln(Rank)×Paid × Low-Ranked		0.104	
		(0.066)	
Year/Month FE	•	•	•
Year/Month FE × Paid	•	•	•
Category FE	•	•	•
Category FE × Paid	•	•	•
Observations	15,958	15,958	28,699
R-squared	0.754	0.754	0.802

Notes: The sample in Columns (1) and (2) are new games (games in their first week on the market). The sample in Column (3) are new non-games (non-game apps in their first week on the market). The outcome variable in both columns are the log of the lower bound of the number of weekly downloads for the apps. "Low-Ranked" apps are apps ranked below 50. Controls include year/month and category fixed effects interacted with a paid app dummy. (Category × Paid) clustered standard errors in parentheses.

B.3 Summary Statistics

Figure B1: Weekly Changes in App Ranking on Top 500 Best-Seller Lists

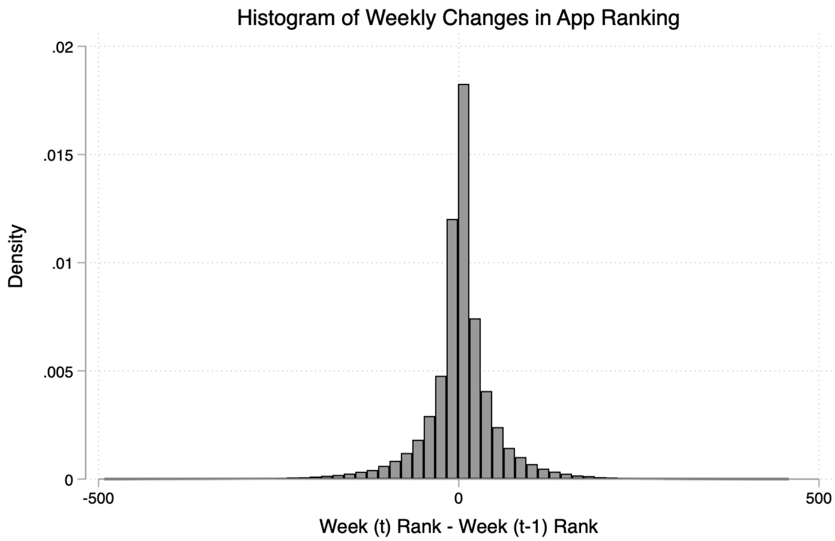


Table B4: **Summary Statistics at the App Type-Month Level**

Variable	Mean	Std. Dev.	N
Game Types			
Number of Apps	7,552	13,386	630
Number of New Apps	720	1,221	630
Non-Game Types			
Number of Apps	33,764	30,960	840
Number of New Apps	2,710	3,026	840

C Appendix C

C.1 Downloads: Average Effects

Table C1: Downloads Difference in Differences Estimates: Average Effects

Outcome Variable:	ln(Tot. Downloads) (1)	ln(Tot. Type Downloads) (2)	ln(App Downloads) (3)	ln(Tot. Type Downloads) (4)	ln(App Downloads) (5)
Games × Post	0.748	1.275	0.162	1.399	0.438
	(0.112)	(0.203)	(0.090)	(0.288)	(0.167)
Games	-7.423				
	(2.860)				
Unit of Observation:	Agg. Game/Non-Game	App-Type	App	App-Type	App
Time Period:	Jan 12/Dec 14	Jan 12/Dec 14	Jan 12/Dec 14	Jan 14/Apr 14	Jan 14/Apr 14
Sample:	All	All	All	All	All
Year/Month FE	•	•	•	•	•
App-Type FE		•		•	
App FE			•		•
App Controls			•		•
Observations	70	1,470	32,964,682	168	5,284,311
R-squared	0.970	0.929	0.953	0.780	0.972

Notes: The sample period in the first three columns is January 2012-December 2014 and in the last two columns is January 2014-April 2014. All app types and apps are considered in each sample. Data in Column (1) consists of monthly observations at the Game/Non-Game level. Data in Columns (2) and (4) consists of monthly observations at the app-type level. Data in Columns (3) and (5) consists of monthly observations at the app level. Outcomes are natural logarithms of downloads at each aggregation level. Controls include year and month fixed effects, game/non-game fixed effects, app-type fixed effects, or app fixed effects, depending on the column. Additional controls in Columns (1)-(3) include game/non-game or app-type specific time trends. Additional app-level controls for Columns (3) and (5) include average app ratings, a dummy for whether the app is free or paid, the price of the app if it is paid, and app age-specific fixed effects. The variable "Games × Post" is a dummy variable equal to 1 for games, or game app-types/apps starting from March 2014. Standard errors are robust to heteroskedasticity in Column (1) and are clustered at the app-type level in the remaining columns.

C.2 Downloads: Category Informativeness Mechanism

After re-categorization, the titles of game categories potentially became more informative about the types of apps present. Before the new categories, consumers looking for music, family or strategy games did not know precisely where to look. After re-categorization, this changed. Consumers had clear information about where different types of apps were located in the store. On the supply side, developers knew that if they produce such a game, there is a clear place for it to be discovered.[8]

I test for this effect in the data. Eight of the eighteen app-types had visibility as game categories before the change: Action and Arcade were grouped together as Arcade & Action, Card and Casino were grouped as Card & Casino. Puzzle was named Brain & Puzzle, and Racing, Sports and Casual games remained unchanged. The remaining app-types did not have pre-existing categories: Adventure, Board, Education, Family, Music, Role Playing, Simulation, Strategy, Trivia and Word games.[9] Consumers should have become much better informed about the location of these app-types after re-categorization and should be able to reach them much faster. If category informativeness plays a role in discovery frictions, downloads for app-types without pre-existing categories should be more affected by re-categorization than app-types with pre-existing categories. I estimate the following regression at the app-type and app level:

$$y_{(j)ct} = \tau^1 \text{Post}_t \times \text{Game}_c + \tau^2 \text{Post}_t \times \text{Game}_c \times \text{No Pre-Existing}_c + \delta_{(j)c} + \delta_t + e_{(j)ct} \quad (1)$$

This regression is estimated using the four months around the re-categorization event (January-April 2014).[10] δ_t and $\delta_{(j)c}$ are month and app-type or app fixed effects, Post_t is a dummy equal to one in March and April 2014, Game_c is a dummy equal to one for all app-types or apps that are games, and No Pre-Existing_c is a dummy equal to one for the ten app-types that did not have categories before March 2014 and zero otherwise. Non-game app-types/apps are the baseline group.[11]

[8] This is especially the case since the new categorization structure already existed on the Apple store for years at that point and developers frequently produce apps for both platforms.

[9] Even though they did not exist in the Google Play store, these were categories in the Apple iOS app store for several years prior.

[10] I also estimate it using only February and March 2014 in Table C3. Results are qualitatively similar but quantitatively smaller.

[11] It is possible that there are some informativeness effects even for app-types with pre-existing categories. For example, consumers searching for Casino apps know after re-categorization that there are only Casino-type apps in the Casino category. Apps belonging to these types also experience changes in the number of other apps in their categories and congestion, which also affects downloads as I show below.

Results from these regressions are in Table C2. They show that app-type-level and app-level downloads increase more for games that did not have pre-existing categories. There is a 44 percent increase in downloads after re-categorization for app-types with pre-existing categories, but the change for an app-types without pre-existing categories is four times as large. The heterogeneity is similar at the app level after controlling for app fixed effects. In Columns (3) and (4), I restrict the sample by excluding some game and non-game types that are very different than game types without pre-existing categories.[12] I still find similar heterogeneity in effects. In Table C4 I also show statistically null effects in response to non-existent events taking place before and after actual re-categorization. These results suggest that re-categorization made the category structure more informative and reduced consumer discovery frictions, increasing downloads.

Table C2: Downloads Difference in Differences Estimates: Category Informativeness

Outcome Variable:	ln(Tot. Type Downloads) (1)	ln(App Downloads) (2)	ln(Tot. Type Downloads) (3)	ln(App Downloads) (4)
Games × Post	0.442	0.220	0.545	0.476
	(0.164)	(0.120)	(0.134)	(0.215)
Games × Post × No Pre-Existing	1.723	1.987	1.635	1.665
	(0.241)	(0.191)	(0.267)	(0.195)
Unit of Observation:	App-Type	App-Type	App	App
Sample Period:	Jan 14/Apr 14	Jan 14/Apr 14	Jan 14/Apr 14	Jan 14/Apr 14
Sample:	All	All	Small Types	Small Types
Year/Month FE	•	•	•	•
App-Type FE	•	•		
App FE			•	•
App Controls			•	•
Observations	168	5,284,311	72	306,956
R-squared	0.916	0.980	0.918	0.935

Notes: The sample period in all columns is January 2014-April 2014. Data in Columns (1) and (3) consists of monthly observations at the app-type level. Data in Columns (2) and (4) consists of monthly observations at the app level. Columns (1) and (2) include all apps. Columns (3) and (4) include all app-types without pre-existing categories and other non-game and game app types with fewer than 20,000 apps in 2012. Outcomes for Columns (1)-(4) are the natural logarithms of downloads at each aggregation level. Additional app-level controls include average app ratings, a dummy for whether an app is free or paid, the price of an app if it is paid and app-age specific fixed effects. The variable "Games × Post" is a dummy variable equal to 1 for games (or game app-types for even columns) during and after March 2014. The variable "Games × Post × No Pre-Existing" is a dummy variable equal to 1 during and after March 2014 only for games/app-types that did not have pre-existing categories before March 2014. Standard errors are clustered at the app-type level.

C.3 Downloads: March and April 2014 Estimates

In the main text, I use January and February 2014 as the "Pre"-policy period for the difference-in-differences download regressions in Tables 3 and C2. March and

[12] Apps without pre-existing categories have fewer apps on average than game types with pre-existing categories or non-game types. I exclude the largest non-game and game types (by the mean number of apps in 2013) to address this concern.

April are the “Post”-policy period. I do this because the policy change happened in the middle of March 2014, so data from that month may not fully reflect the policy change.

Table C3 replicates key regressions from Tables 3 and C2 from the main text using only data from February and March 2014. The “Pre” policy period is February 2014 and the “Post” policy period is March 2014. Results are qualitatively similar but quantitatively smaller than in the main text.

Table C3: Downloads Difference in Differences: February and March Data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcome Variable:	ln(Tot Type Dwnlds)	ln(Downloads)	ln(Tot Type Dwnlds)	ln(Downloads)	ln(Tot Type Dwnlds)	ln(Downloads)	Post/Pre Δ ln(Downloads)
Games × Post	1.081	0.286	0.251	0.084	0.144	0.072	
	(0.220)	(0.132)	(0.113)	(0.050)	(0.012)	(0.009)	
Games × Post × No Pre-Existing			1.494	1.791			
			(0.242)	(0.082)			
Games × Post × Small Type					0.632	0.921	
					(0.010)	(0.119)	
Post/Pre Δ ln(N Apps)							-0.594
							(0.032)
Unit of Observation	App-Type	App	App-Type	App	App-Type	App	App
Sample Period	Feb 14/Mar 14	Feb 14/Mar 14	Feb 14/Mar 14	Feb 14/Mar 14	Feb 14/Mar 14	Feb 14/Mar 14	Feb 14/Mar 14
Sample	All	All	All	All	All Non-Games + Card, Casino, Arcade and Action	All Non-Games + Card, Casino, Arcade and Action	All Games
Observations	84	2,574,302	84	2,574,302	56	2,330,302	142,419
R-squared	0.833	0.981	0.944	0.989	0.998	0.992	0.868
Year/Month FE	•	•	•	•	•	•	
App-Type FE	•		•		•		
App FE		•		•		•	

Notes: The sample period in all columns is February and March 2014. Data in Columns (1), (3) and (5) consists of monthly observations at the app-type level. Data in Columns (2), (4), (6) and (7) consists of monthly observations at the app level. Columns (1)-(4) include all apps. Columns (5) and (6) include all non-game apps and Arcade, Action, Card and Casino game apps. Column (7) includes all game apps. Outcomes for Columns (1)-(6) are the natural logarithms of downloads at each aggregation level. The outcome for Column (7) is the difference between the natural log of app downloads in March 2014 and downloads in February 2014. Controls include year and month fixed effects and app-type or app fixed effects. Column (7) does not have fixed effects because it is a cross sectional regression in first-differences. Additional app-level controls include average app ratings, a dummy for whether an app is free or paid, the price of paid apps and app-age specific fixed effects. The variable “Games × Post” is a dummy variable equal to 1 for games, or game app-types in March 2014. “No Pre-Existing” is a dummy variable equal to 1 for apps or app-types with no pre-existing categories before March 2014. “Small Split” is a dummy variable equal to 1 for Action and Casino games. “Post/Pre ln(N Apps in Category)” is the difference in the natural log of the number of apps in the category of app j in March 2014 and the number of apps in February 2014. Standard errors are clustered at the app-type level for Columns (1)-(6) and are robust to heteroskedasticity in Column (7).

C.4 Downloads: Placebos, Paid Apps Only and Alternative Outcomes

C.4.1 Downloads: Placebo Time Periods

Most of the main results in Section III are computed using a restricted data sample of four months, comparing January and February 2014 to March and April 2014. A possible concern may be that estimated effects are not caused by re-categorization but by diverging time-trends between games and non-games or across game types. To test whether this is the case, I re-estimate the regressions using two comparable time periods without a re-categorization event: November 2013 to February 2014, and March 2014 to June 2014. For each sample, I estimate the effects of a non-existent re-categorization event: between December and January for the first sample, and between April and May for the second sample.

The first "placebo" sample verifies that download trends between games and non-games or across game types were not diverging before re-categorization took place. It also helps test whether it was the actual re-categorization or the *announcement* of re-categorization in December 2013 changed download outcomes. If changes in downloads were actually caused by re-categorization and improved consumer discovery, there is no reason to expect statistically significant differences in downloads before. The second "placebo" sample further verifies the effects of re-categorization. The policy was a permanent event - a consumer in May 2014 should have had as easy a time finding the "Music" game category as a consumer in April 2014. If re-categorization improved consumer discovery technology, these improvements should be locally persistent over time.[13]

I show results using alternative time periods in Table C4. These estimates replicate the main specifications shown in Tables 3 and C2.[14] Panel (a) shows results using the Nov 2013 - Feb 2014 sample and panel (b) shows results using the March 2014 - June 2014 sample.

Nearly all estimates for the alternative time periods are statistically null. They are also generally substantially smaller in magnitude than estimates in the main text. There is some evidence of heterogeneous time-trends across game types without pre-existing categories and game types with pre-existing categories prior to re-categorization in Columns (3) and (4) in panel (a). However, relative to the baseline group of non-game types or apps, the total change in downloads for game app-types without pre-existing categories is still null. For both Columns (3) and (4), the sum of the "Games × Placebo Post" and "Games × Placebo Post × No Pre-Existing" is not statistically significantly different from zero. These results show that differential changes in downloads between games and non-games occurred only during re-categorization. These results also show that the effects of re-categorization

[13]Over a longer period of time, changes in product assortment and entry may introduce additional congestion costs into the market, mitigating some immediate decreases in discovery costs. The changes in category informativeness, however, should be very persistent over time.

[14]Full time-varying estimates of treatment effects for specifications where I use data from January 2012 to December 2014 are in Figure C1. They also show the main download effects appear only following the actual re-categorization.

on downloads are persistent. Downloads four months after re-categorization were not statistically significantly different than the month after re-categorization. This suggests that the re-categorization event directly caused the change in downloads, consistent with permanent reductions in search costs for consumers due to an improvement in category informativeness and a reduction in the number of apps per category.

Table C4: Downloads Difference in Differences: Alternative Time Periods

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome Variable:	ln(Type Downloads)	ln(Downloads)	ln(Type Downloads)	ln(Downloads)	ln(Type Downloads)	ln(Downloads)
			Panel (a): Nov 2013 - Feb 2014			
Games × Placebo Post	-0.131	-0.291	-0.219	-0.302	-0.210	-0.312
	(0.137)	(0.199)	(0.154)	(0.198)	(0.132)	(0.149)
Games × Placebo Post × No Pre-Existing			0.158	0.121		
			(0.027)	(0.035)		
Games × Placebo Post × Small Type					0.019	0.105
					(0.039)	(0.111)
Observations	168	4,892,297	168	4,892,297	112	4,427,584
R-squared	0.987	0.978	0.987	0.978	0.968	0.982
			Panel (b): March 2014 - June 2014			
Games × Placebo Post	0.223	0.065	0.110	0.056	0.124	0.079
	(0.202)	(0.140)	(0.120)	(0.117)	(0.107)	(0.106)
Games × Placebo Post × No Pre-Existing			0.204	0.064		
			(0.147)	(0.168)		
Games × Placebo Post × Small Type					0.082	-0.030
					(0.074)	(0.119)
Observations	168	5,496,539	168	5,496,539	112	4,935,476
R-squared	0.900	0.985	0.909	0.985	0.955	0.986
Unit of Observation:	App-Type	App	App-Type	App	App-Type	App
Sample	All	All	All	All	All Non-Games + Card, Casino, Arcade and Action	All Non-Games + Card, Casino, Arcade and Action
Year/Month FE	•	•	•	•	•	•
App-Type FE	•		•		•	
App FE		•		•		•

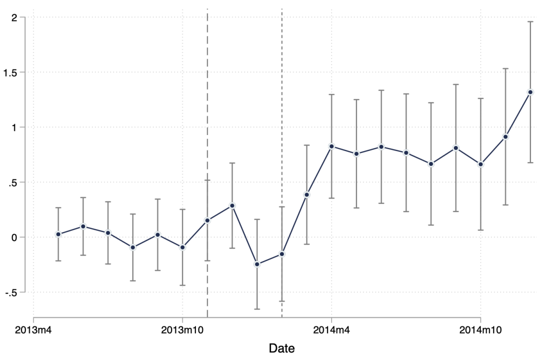
Notes: The sample period in panel (a) covers November 2013 to February 2014. The sample period in panel (b) covers March 2014 to June 2014. Sample in Columns (1), (3) and (5) consists of monthly observations at the app-type level. Sample in Columns (2), (4) and (6) consists of monthly observations at the app level. Outcomes are natural logarithms of downloads at each aggregation level. Controls include year and month fixed effects, app-type fixed effects, or app fixed effects, depending on the column. Additional app-level controls for Columns (2), (4) and (6) include average app ratings, a dummy for whether the app is free or paid, the price of the app if it is paid, and app age-specific fixed effects. The variable "Games × Placebo Post" is a dummy variable equal to 1 for games, or game app-types during and after January 2014 for panel (a) and during and after May 2014 for panel (b). Standard errors are clustered at the app-type level.

C.4.2 Downloads: Paid Apps

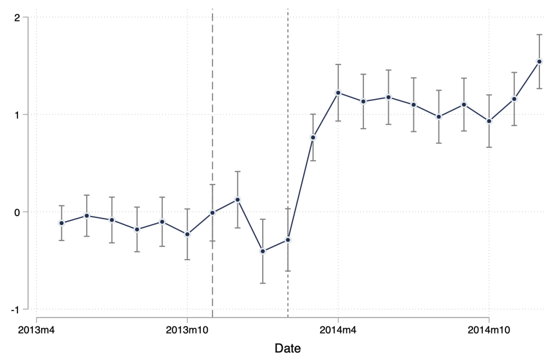
Downloads of free apps may not be accurate proxies of consumer app usage, as consumers can easily install and uninstall such apps from their phones without fully inspecting or using them. This is not the case for paid apps that consumers are required to spend money on upfront. I re-estimate the main regressions from Tables

Figure C1: Downloads Timing Tests

(a) Outcome: ln(Total Downloads)



(b) Outcome: ln(Total Type Downloads)



Notes: Each panel shows estimates of coefficients τ_t from Equation 3 at different aggregation levels. Panel (a) is estimated at the game/non-game level and panels (b) is estimated at the app-type level. Data from January 2012 to December 2014 is used throughout. Additional controls in each regression include year/month fixed effects, game/non-game or app-type fixed effects, and game/non-game or app-type specific trends. Standard errors for panel (a) are robust to heteroskedasticity and standard errors for panel (b) are clustered at the app-type level. 95% confidence intervals shown. In each panel, the first dashed vertical line represents the announcement of re-categorization and the second dashed vertical line represents the start of the re-categorization period.

3 and C2 in the main text after restricting the sample to paid apps. I also test for placebo effects using alternative time periods, as in Table C4.

Estimates for the sample of paid apps are in Table C5. Odd columns show estimates from aggregated app-type regressions with app-type fixed effects. Even columns show estimates from app-level regressions and even columns are at the app level with app fixed effects. Panel (a) shows estimates of the regressions for the January 2014 - April 2014 period, testing for the effects of a re-categorization event in March. Panel (b) shows estimates using the November 2013 - February 2014 period, testing the effects of a non-existent re-categorization event between December and January. Panel (c) similarly shows estimates using the March 2014 - June 2014 period, with a non-existent re-categorization event between April and May.

In each panel, the first two columns show the baseline average effects. The next two columns test for heterogeneity across game types that had pre-existing categories before the policy and those that did not. Such heterogeneity identifies changes in discovery costs through increasing informativeness. The last two columns test for heterogeneity across Arcade, Action, Card and Casino game types, where Action and Casino were much smaller before re-categorization. Such heterogeneity should identify changes in discovery costs through reducing the number of apps per category and congestion.

Results are consistent with those in the main text and the robustness checks above. Panel (a) shows that on average, downloads for paid games increased over non-games after re-categorization. Paid games belonging to types without pre-existing categories and games belonging to types with fewer apps are driving the main effects. Estimates are statistically significant at the 95% confidence level and are larger than in the main text. Results for the two "placebo" events, before and after the actual re-categorization, show statistically null effects. These estimates confirm that consumer discovery costs fell in response to re-categorization.

C.4.3 Downloads: Alternative Outcomes

I use three alternative outcome variables to test the robustness of estimates in Section III. Two of the alternative outcomes do not rely on the procedure described in Appendix B.2. As discussed by Liebowitz and Zentner (2020), the parametric assumptions used to generate most monthly download values in Appendix B.2 can produce biased estimates of actual downloads.

The first alternative outcome restricts the sample of apps to new apps: apps that entered the store in month t. For these apps, the approximation bias is minimal, since they are used to fit the model in Appendix B.2.

The second alternative outcome is a simpler proxy for monthly downloads: the difference in the number of user ratings for an app between two periods. For app j, downloads for period t are approximated by the number of user ratings in period t minus the number of user ratings in period $t-1$ ($\text{N Ratings}_{jt} - \text{N Ratings}_{jt-1}$).[15] This proxy relies on a simple, intuitive relationship - if a certain proportion of users who download an app also rate it, apps with more downloads will also have more ratings. Ratings on Google Play have to come from downloads, and it is unlikely that users will wait over a month to rate an app they downloaded. Such proxies have been used previously in papers studying mobile apps, such as Kummer and Schulte (2019). This approach has limitations, as the relationship between downloading and rating is not necessarily strictly monotonic in the number of downloads. Some types of apps may be very frequently downloaded but not frequently rated, whereas other apps are both frequently downloaded and rated. The proportion of users who rate apps may also decrease in app popularity. This would create a bias in the download

[15]On occasion, ratings and reviews disappear from the Google Play Store for various reasons including service term violations and the number of ratings falls between period $t-1$ and period t. This occurs for less than 0.7% of observations. I bound changes in ratings to zero from below.

Table C5: Downloads Difference-in-Differences: Paid App Sample

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome Variable:	ln(Type Dwnlds)	ln(Dwnlds)	ln(Type Dwnlds)	ln(Dwnlds)	ln(Type Dwnlds)	ln(Dwnlds)
			Panel (a): Jan 2014 - Apr 2014			
Games × Post	1.536	0.489	0.594	0.252	0.332	0.255
	(0.228)	(0.102)	(0.163)	(0.040)	(0.069)	(0.020)
Games × Post × No Pre-Existing			1.697	2.178		
			(0.315)	(0.220)		
Games × Post × Small Type					1.322	1.560
					(0.237)	(0.154)
Observations	168	972,440	168	972,440	112	883,472
R-squared	0.761	0.978	0.876	0.986	0.953	0.990
			Panel (b): Nov 2013 - Feb 2014			
Games × Placebo Post	0.052	-0.024	0.064	-0.029	-0.038	-0.022
	(0.034)	(0.057)	(0.043)	(0.057)	(0.056)	(0.045)
Games × Placebo Post × No Pre-Existing			-0.022	0.049		
			(0.065)	(0.019)		
Games × Placebo Post × Small Type					0.314	0.020
					(0.126)	(0.043)
Observations	168	942,428	168	942,428	112	858,336
R-squared	0.990	0.989	0.990	0.989	0.989	0.989
			Panel (c): Mar 2014 - Jun 2014			
Games × Placebo Post	0.255	0.103	0.142	0.090	0.142	0.093
	(0.138)	(0.069)	(0.084)	(0.072)	(0.084)	(0.052)
Games × Placebo Post × No Pre-Existing			0.203	0.096		
			(0.162)	(0.196)		
Games × Placebo Post × Small Type					0.104	0.023
					(0.045)	(0.101)
Observations	168	958,186	168	958,186	112	866,545
R-squared	0.925	0.990	0.934	0.990	0.963	0.990
Unit of Observation:	App-Type	App	App-Type	App	App-Type	App
Sample	All Paid	All Paid	All Paid	All Paid	All Paid Non-Games + Paid Card, Casino, Arcade and Action	All Paid Non-Games + Paid Card, Casino, Arcade and Action
Year/Month FE	•	•	•	•	•	•
App-Type FE	•		•		•	
App FE		•		•		•

Notes: The sample throughout all panels and columns only includes *paid* apps with non-zero prices. The sample period in panel (a) covers January 2014 to April 2014. The sample period in panel (b) covers November 2013 to February 2014. The sample period in panel (c) covers March 2014 to June 2014. Sample in Columns (1), (3) and (5) consists of monthly observations at the app-type level. Sample in Columns (2), (4) and (6) consists of monthly observations at the app level. Outcomes are natural logarithms of downloads at each aggregation level. Controls include year and month fixed effects, app-type fixed effects, or app fixed effects, depending on the column. Additional app-level controls for Columns (2), (4) and (6) include average app ratings, the price of the app, and app age-specific fixed effects.The variable "Games × Post" is a dummy variable equal to 1 for games, or game app-types during and after March 2014 for panel (a). The variable "Games × Placebo Post" is a dummy variable equal to 1 for games, or game app-types after during and after January 2014 for panel (b) and during and after May 2014 for panel (c). Standard errors are clustered at the app-type level.

proxy. For this reason, I choose to use downloads calculated according to B.2 as the main specification.

The last alternative outcome is the absolute number of predicted downloads rather than the natural log of downloads (using the predicted measure of downloads from Online Appendix B.2).

Results using these three outcomes are in Table C6. There are four panels in the table. For ease of comparison, Panel (a) provides results using the baseline outcome from the main text. Panel (b) shows results using only new apps. Panel (c) shows results using the difference in the number of ratings to approximate downloads. Panel (d) shows results using the absolute number of predicted downloads. Odd columns aggregate data at the app-type level, and even columns use app level data. App level fixed effects are included for regressions in panels (a), (c) and (d). Panel (b) only includes app-type fixed effects, as I only observe each new app once. App level regressions in each panel are done using all non-game apps and game apps belonging to app-types without categories before the policies. I pick this sample as discovery costs should fall for this set of game apps (see panel (a) of Table C2). Columns (1) and (2) use the baseline January 2014 - April 2014 four-month sample period. Columns (3) and (4) use November 2013 to February 2014 as the sample period, with a "placebo" event between December 2013 and January 2014. Columns (5) and (6) use March 2014 to June 2014 as the sample period, with a "placebo" event between April 2014 and May 2014.

Results are qualitatively equivalent to the main specification. Results for ratings based downloads are different in magnitude than in the main text because of the different definition. However, downloads statistically significantly increase for games relative to non-games in March and April 2014 relative to January and February. The same does not happen in May and June 2014 relative to March and April, or in January and February 2014 relative to November and December 2013. Estimates in panel (b) using the sample of new apps are also similar to those in the main text.

Table C6: Downloads Difference in Differences Estimates: Alternative Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel (a): Baseline			
Outcome Variable:	ln(Tot. Type Downloads)	ln(App Downloads)	ln(Tot. Type Downloads)	ln(App Downloads)	ln(Tot. Type Downloads)	ln(App Downloads)
Games × Post	1.399	2.214				
	(0.288)	(0.302)				
Games × Placebo Post			-0.131	-0.174		
			(0.137)	(0.173)		
Games × Placebo Post					0.223	0.120
					(0.202)	(0.273)
Observations	168	4,646,394	168	4,294,910	168	4,825,839
R-squared	0.780	0.982	0.987	0.982	0.900	0.986
			Panel (b): New App Sample			
Outcome Variable:	ln(Tot. New Type Downloads)	ln(New App Downloads)	ln(Tot. New Type Downloads)	ln(New App Downloads)	ln(Tot. New Type Downloads)	ln(New App Downloads)
Games × Post	1.840	2.440				
	(0.318)	(0.128)				
Games × Placebo Post			0.486	-0.315		
			(0.207)	(0.172)		
Games × Placebo Post					-0.045	-0.087
					(0.085)	(0.120)
Observations	168	378,987	168	481,961	168	363,619
R-squared	0.821	0.622	0.849	0.507	0.931	0.706
			Panel (c): Ratings Based Downloads			
Outcome Variable:	ln(Tot. Type Δ Ratings)	ln(App Δ Ratings)	ln(Tot. Type Δ Ratings)	ln(App Δ Ratings)	ln(Tot. Type Δ Ratings)	ln(App Δ Ratings)
Games × Post	0.092	0.196				
	(0.025)	(0.062)				
Games × Placebo Post			0.068	0.172		
			(0.032)	(0.147)		
Games × Placebo Post					0.072	0.050
					(0.039)	(0.062)
Observations	168	4,646,680	168	4,284,464	168	4,829,472
R-squared	0.992	0.918	0.987	0.877	0.991	0.917
			Panel (d): Absolute Downloads			
Outcome Variable:	Tot. Type Downloads	App Downloads	Tot. Type Downloads	App Downloads	Tot. Type Downloads	App Downloads
Games × Post	1301604.410	739.440				
	(459,123.970)	(227.632)				
Games × Placebo Post			-217,949.749	-198.974		
			(313,913.912)	(106.365)		
Games × Placebo Post					417,688.494	110.491
					(383,730.890)	(184.156)
Observations	168	4,646,394	168	4,294,910	168	4,825,839
R-squared	0.854	0.780	0.843	0.817	0.908	0.931
Unit of Observation:	App-Type	App	App-Type	App	App-Type	App
Sample Period:	Jan 14/Apr 14	Jan 14/Apr 14	Nov 13/Feb 14	Nov 13/Feb 14	Mar 14/Jun 14	Mar 14/Jun 14
Sample	All	All Non-Games + Games w/o Pre-Exist Cats.	All	All Non-Games + Games w/o Pre-Exist Cats.	All	All Non-Games + Games w/o Pre-Exist Cats.
Year/Month FE	•	•	•	•	•	•
App-Type FE	•		•		•	
App FE		•		•		•

Notes: Sample for odd columns includes all apps and for even columns includes all non-games and Adventure, Board, Education, Family, Music, Role Playing, Simulation, Strategy, Trivia and Word games. Sample period in Cols (1)-(2) covers Jan/Apr 2014. Sample period in Cols (3)-(4) covers Nov 2013 - Feb 2014. Sample period in Cols (5)-(6) covers Mar/June 2014. Odd columns sample Outcomes are defined as the title of each column/panel combination. Controls include year and month fixed effects, app-type fixed effects, or app fixed effects, depending on the column. Additional app-level controls for even columns include average app ratings, a dummy for whether the app is free or paid, the price of the app if it is paid, and app age-specific fixed effects. Panel (b) does not include app-level fixed effects. In Cols (1)-(2) "Games × Post" is a dummy variable equal to 1 for games, or game app-types during and after March 2014. In Cols (3)-(4) "Games × Placebo Post" is a dummy variable equal to 1 for games, or game app-types during and after January 2014. In Cols (5)-(6) "Games × Placebo Post" is a dummy variable equal to 1 for games, or game app-types during and after May 2014. Standard errors are clustered at the app-type level.

C.5 Downloads: Changes in Number of Apps per Category

This section shows how, for a given app, the number of other apps in its category changes between November/December 2013 and January/February 2014, January/February 2014 and March/April 2014, and March/April 2014 and May/June 2014. I do this by first calculating, for each app, the average number of other apps in its category in each two-month period.[16] Then I calculate, for each app, the difference between two successive periods.

The distribution of changes appears in Figure C2. This figure has three panels representing the three sets of changes I examine. The distribution of changes in panels (a) and (c) is entirely different from the distribution of changes in panel (b). In panels (a) and (c), the number of other apps in the category of an app increase. This is consistent with the general growth in the number of apps on Google Play over time. Panel (b) shows that between Jan/Feb and Mar/Apr 2014, all apps experienced a drop in the number of other apps in their category. The variation in this drop represents differences between what category the app belonged to before re-categorization and its category/app-type after re-categorization.

C.6 Downloads and Long Run Entry

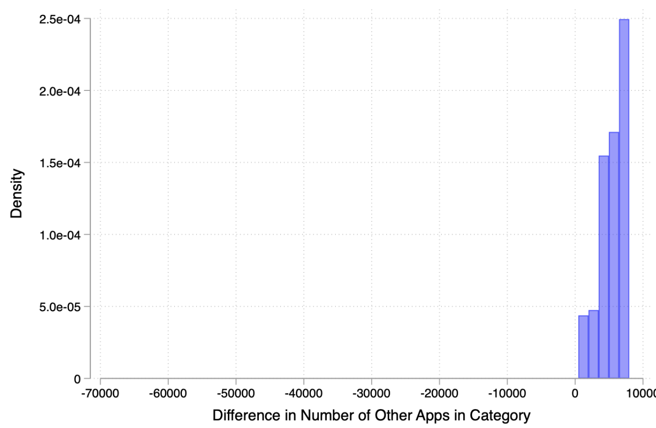
In Column (3) of Table 3 in the main text, I show evidence of the effects of short-run changes in the number of apps in an app's category on app-level downloads. I do this using short-run changes induced by re-categorization, which move apps from broad categories to narrow categories that reflect their app types. Since where the apps end up is primarily determined by their pre-existing app-type and entry does not change much, this primarily reflects congestion rather than changes in the competition intensity (as reflected by the number of substitutes) for each app.

However, this does not necessarily mean that longer run changes in the number of apps of each app-type will have similar effects. As mentioned above, increases in the number of apps of each app-type could affect both competition intensity and congestion. I test for this in the data by relating long run differences in app-type entry to long run differences in downloads in the post re-categorization period. For a given app j, I calculate the difference in their downloads between December 2014 and March 2014. I then regress this difference on the difference in the number of apps of their type (which also coincides with their category) between December 2014

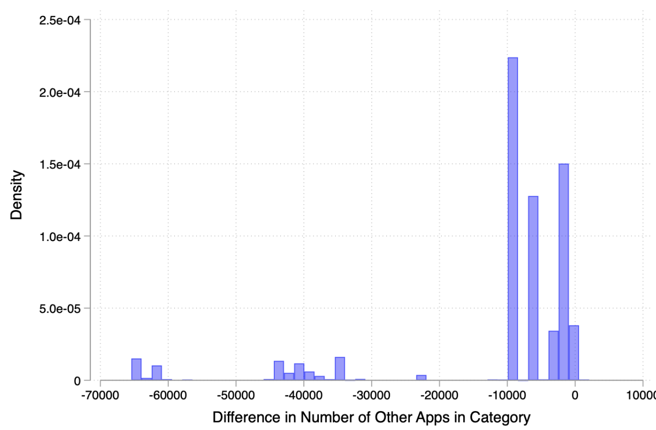
[16] I refer to categories here rather than app types since consumers use the stated category structure to search. See Section I for additional discussion. I use two-month periods, since these are comparable to the time periods I use in Table 3.

Figure C2

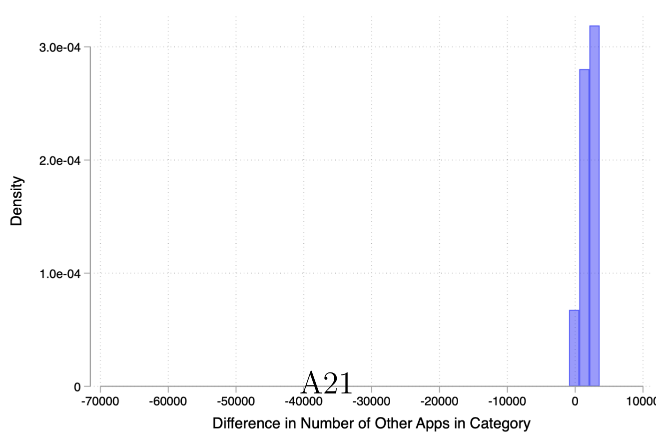
(a) ln(N Apps Jan/Feb 2014) - ln(N Apps Nov/Dec 2013)



(b) ln(N Apps Mar/Apr 2014) - ln(N Apps Jan/Feb 2014)



(c) ln(N Apps May/Jun 2014) - ln(N Apps Mar/Apr 2014)



Notes: Each panel shows the distribution of changes in app-level changes in the number of other apps in their category over time. For each app j, I calculate the difference in the natural log of the number of apps in their category between two successive periods. If app j is in category c^* in period t and category d^* in period $t+1$, the difference is $ln(N_{d^*,t+1}) - ln(N_{c^*,t})$. In panel (a), the difference is between Jan/ Feb 2014 (on average) Nov/ Dec 2013. In panel (b), the difference is between Jan/Feb 2014 and Mar/Apr 2014. In panel (c), the difference is between Mar/Apr 2014 and May/Jun 2014.

and March 2014. The estimating equation is as follows:

$$ln(Downloads_{j,Dec}) - ln(Downloads_{j,Mar}) = \alpha(ln(NApps)_{j,Dec}) - ln(NApps)_{j,Mar}) + \beta X_j + \epsilon_j \quad (2)$$

where X_j are app characteristics I control for to account for unobservable heterogeneity not fully absorbed by the within-app differencing.

Estimates of this regression are in Table C7. The coefficient on the difference in the number of apps in the category is -0.66, suggesting that a one percent increase in the number of apps reduces app downloads by 0.66%.

Table C7: Long Run Changes in Downloads and Entry

	(1)
Outcome:	Dec/Mar Δ ln(Downloads)
Dec/Mar Δ ln(N Apps)	-0.655
	(0.016)
Unit of Observation	App
Sample	All Games
Sample Period	Mar and Dec 2014
App Controls	•
Observations	121,134
R-squared	0.556

Notes: Sample includes all apps present in both March and December 2014. App-level controls include average app ratings, a dummy for whether the app is free or paid, the price of the app if it is paid, and app age-specific fixed effects. Standard errors are robust to heteroskedasticity.

These long-run estimates of the elasticity between changes in the number of apps in a category and app-level downloads are strikingly similar to short-run estimates in the main text. The coefficient on the short run re-categorization driven changes in the number of apps on the number of downloads is -0.65. This is reassuring, since the short-run re-categorization effect on congestion I identify in the main text seems to also be operating in the longer run. If changes in longer run entry were also generating pressure on app downloads through competition from additional substitutes, I would expect the coefficient in the long run regression to be substantially larger (in absolute terms). One possibility is that competition is already intense in the app market in March 2014 such that additional entry between March and December does not significantly increase it.

The OLS regressions estimated in this section may be subject to endogeneity concerns, as both downloads and app-type entry could be determined by common unobservable shocks. However, in many ways, the evolution of app-type entry after re-categorization is driven by changes caused by re-categorization. In Section C.7.1 I show that changes in entry after re-categorization are driven by whether the app-type's discovery costs fell during re-categorization. This means that app entry in

November 2014 (e.g., between November and December) was largely driven by the re-categorization which happened eight months prior.

C.7 Entry

C.7.1 Entry: Mechanisms

Sections III and Appendix C.2 show that re-categorization produced two main demand-side effects: an increase in the informativeness of categories in the store, and a reduction in congestion. Both of these improved consumers' ability to effectively browse the store, but some app-types (and ex-post categories) were more affected than others. The supply-side entry effects shown in Section IV should be driven by these demand-side mechanisms. In that case, entry effects should display similar heterogeneity to download effects. For example, consumers became more informed about app-types that did not feature in the pre-policy categories. Entry in these app-types should increase more as well. I test for such heterogeneity in Table C8. As in Table C2, Column (1) tests for the effects of changes in the informativeness of categories by comparing app-types with and without pre-existing categories. As in Table 3, Column (2) tests for the effects of changes in congestion costs by comparing small and large app-types among those split from two pre-policy game categories. Small app-types had fewer apps before re-categorization, and they experience greater decreases in congestion after re-categorization.

Estimates confirm that average effects in Table 5 were primarily driven by app-types with greater changes in discovery costs. Entry increased more for game app-types that did not have pre-existing categories, as compared to game app-types with pre-existing categories (relative to non-game types). Among the app-types split off from pre-policy game categories, smaller app-types also had bigger changes in entry. App-types whose discovery costs were less affected by the policy have statistically null changes for both of the main supply-side outcomes. These results are also robust to alternative outcomes, such as the absolute number of entrants (Table C9).

Results in Tables 5 and C8 suggest that changes in consumer discovery costs are the main driving mechanism for product assortment changes in this market. The heterogeneity in entry effects also reflects theoretical predictions about the effects of discovery cost changes for different product types. Bar-Isaac et al. (2012) predict that "niche" products that benefit more from search cost reductions will experience the greatest increase in assortment. In this setting, the definition of "niche" products can include either app-types that had no pre-existing categories or "small types" that were relatively marginalized under the initial category structure.

Table C8: Entry Difference in Differences Estimates: Discovery Cost Channels

Outcome Variable:	ln(N Entrants) (1)	ln(N Entrants) (2)
Games × Post	0.251	-0.086
	(0.137)	(0.069)
Games × Post × No Pre-Existing	0.550	
	(0.138)	
Games × Post × Small Type		0.833
		(0.202)
Unit of Observation	App-Type	App-Type
Time Period	Jan 12 / Dec 14	Jan 12 / Dec 14
Sample	All	All Non-Games + Action, Arcade Card and Casino
Year/Month FE	•	•
App-Type FE	•	•
Observations	1,470	980
R-squared	0.976	0.959

Notes: The sample period in all columns is January 2012-December 2014. Data in all columns includes monthly observations at the app-type level. Sample in Column (1) includes all game and non-game app-types. Sample in Column (2) includes all non-game app-types and Arcade, Action, Card and Casino game types. Outcomes in both columns are the natural log of the number of entrants in each app-type. Controls include year and month fixed effects, app-type fixed effects and app-type specific time trends. The variable "Games × Post" is a dummy variable equal to 1 for games (or game app-types for even columns) during and after March 2014. The variable "Games × Post × No Pre-Existing" is equal to 1 during and after March 2014 only for app-types that did not have pre-existing categories before March 2014. The variable "Games × Post × Small Type" is a dummy equal to 1 during and after March 2014 only for Action and Casino game app-types. Standard errors in all columns are clustered at the app-type level.

C.7.2 Entry: Alternative Outcomes

Table C9 replicates Table 5 in the main text with alternative outcome variables. It uses *absolute* entry numbers. These regressions confirm the results from the log-transformed estimates in the main text. Absolute entry for the average app-type increases by about 1,500 apps, and for all games by almost 35,000 apps after re-categorization. As in the log estimates in Table 5, these are large treatment effects, given the size of the average game app-type. Columns (3) and (4) show that app-types where discovery costs fall by more in response to re-categorization are the ones driving the effects.

Table C9: Entry Difference in Differences Estimates with Alternative Outcomes

	Panel (a): Absolute Number of Entrants			
Outcome:	N Entrants	N Type Entrants	N Type Entrants	N Type Entrants
Games × Post	34,658.448	1,463.367	1,134.432	854.808
	(16,259.260)	(310.944)	(394.645)	(348.410)
Games × Post × No Pre-Existing			592.083	
			(296.673)	
Games × Post × Small Type				1,214.406
				(438.110)
Observations	70	1,470	1,470	980
R-squared	0.876	0.775	0.776	0.776
Unit of Observation	Game/Non-Game	App-Type	App-Type	App-Type
Time Period	Jan 12 / Dec 14	Jan 12 / Dec 14	Jan 12 / Dec 14	Jan 12 / Dec 14
Sample	All	All	All	All Non-Games + Action, Arcade Card and Casino
Year/Month FE	•	•	•	•
App-Type FE		•		•

Notes: Sample period in all columns and panels is January 2012 to December 2014. Sample in Column (1) includes monthly observations at the Game/Non-Game level. Sample in Columns (2) and (3) includes all game and non-game observations at the app-type level. Sample in Column (4) includes all non-game and Action, Arcade, Card and Casino app-type observations at the monthly level. Outcomes are the absolute number of new entrants at the game/non-game or app-type level. Outcomes in panel (b) are the average share of 1-star ratings at the game/non-game or app-type level. Controls include year/month fixed effects, a "Game" category group dummy for odd columns and app type fixed effects. Additional controls include game/non-game or app-type specific time trends. The variable "Games × Post" is equal to 1 for games (or game types for even columns) during and after March 2014 and zero otherwise. Standard errors are robust to heteroskedasticity in Column (1) and clustered at the app-type level otherwise.

C.7.3 Entry: Timing Tests

I allow treatment effects to vary over time by introducing interactions between monthly date dummies and the treatment group dummy. I estimate timing tests both at the aggregate game/non-game level and at the less aggregate app-type level. The estimating equation, for app-type c at time t is:

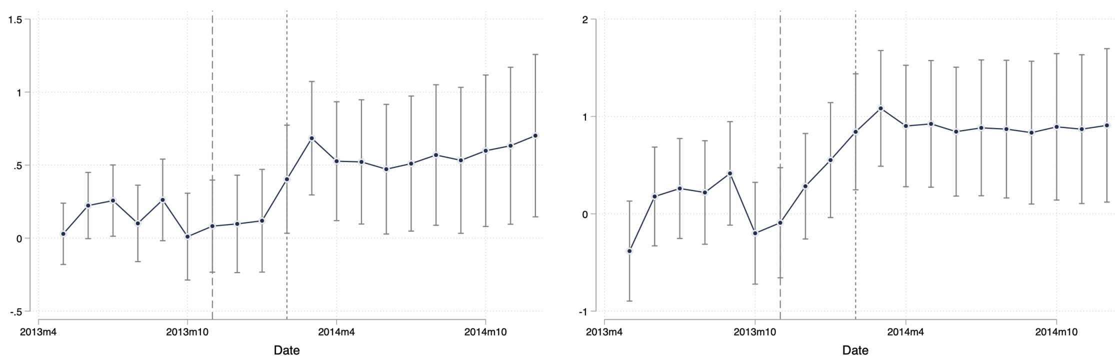
$$y_{ct} = \sum_{t=\text{re-cat. month}-10}^{\text{re-cat. month}+10} \tau_t \; (Game_c \times D_t) + \delta_c + \delta_t + \epsilon_{ct} \tag{3}$$

where y_{ct} is an outcome, and where τ_ts now capture period specific treatment effects relative to a baseline period. D_t is a dummy equal to one for observations during month t and zero otherwise. Since I have 10 periods after re-categorization, I test the 10 periods before re-categorization for parallel pre-trends relative to the time before April 2013. As before, I include game/non-game, app-type and time fixed effects, and game/non-game or app-type specific time trends. Figure C3 shows the period specific treatment effects for the main measure of entry used in the main text. For each outcome, game/non-game level results are on the left-hand panel, and app-type results are on the right-hand panel.

Entry estimates show that treatment effects become statistically significantly different from zero exactly around re-categorization. There are no treatment effect

Figure C3: Entry Timing Tests

(a) Outcome: ln(N Entrants) (b) Outcome: ln(N Type Entrants)



Notes: Each panel shows estimates of coefficients τ_t from Equation 3 at different aggregation levels and with different outcomes. Panel (a) is estimated at the game/non-game level. Panels (b) is estimated at the app-type level. Data from January 2012 to December 2014 is used throughout. Additional controls in each regression include year/month fixed effects, game/non-game or app-type fixed effects, and game/non-game or app-type specific trends. Standard errors for panel (a) are robust to heteroskedasticity and standard errors for panel (b) are clustered at the app-type level. 95% confidence intervals shown. In each panel, the first dashed vertical line represents the announcement of re-categorization and the second dashed vertical line represents the start of the re-categorization period.

estimates which are statistically significantly different than zero (at the 95 percent confidence level) in the 10 periods before February 2014. February 2014 itself (two months following the announcement) has a statistically significant positive coefficient, likely representing a response by developers to the announcement of new categories in December 2013. Some apps may have entered the market early to position themselves in anticipation of the change.[17] Entry response happens quickly after the announcement since apps have short development time. Developers can create simple apps in as little as a month.[18] Point estimates are highest right after re-categorization takes place.

[17]The announcement did not set a strict date for the implementation of new categories, but said that the change will happen in the first quarter of 2014 (9to5Google.com)

[18]New entry could have come from multiple sources. Developers creating completely new apps, porting existing apps from the iOS store, or releasing already developed products into the market early.

C.8 Prices

Figure C4 plots three graphs showing price patterns in the Google Play Store. Panel (a) shows the ratio of mean paid game prices over mean paid non-game prices. Panel (b) shows the ratio of mean *new* paid game prices over mean *new* paid non-game prices. Prices do not appear to change substantially.

Panel (a) shows average prices for all games falling as compared to non-games, potentially due to increasing importance of in-app advertising and in-app purchases in the app economy. After re-categorization, average game prices increase and the ratio of game to non-game prices stabilizes. Lower discovery costs from the re-categorization could be the cause of the price changes.[19] However, in absolute terms, the magnitudes of changes are small. In panel (b), it is apparent than there are no substantial differences in the prices of new games relative to non-games. The price ratio before and after is similar on average.

I estimate difference-in-differences regressions with both average prices and average entrant prices as outcome variables. Results for these regressions are in Table C10 at the game/non-game and at the app type. They show that there are no statistically significant differences between game and non-game paid app prices after re-categorization as compared to before. There is also no statistically significant heterogeneity (at the 95% confidence level) across game app-types that were more or less affected by re-categorization. This is true regardless of whether I look at all paid apps in panel (a) or only at new paid apps in panel (b).

Panel (c) of Figure C4 shows the ratio of the share of new paid games appearing in a given month (as a percentage of the total number of new games), over the share of new paid non-games (as a percentage of the total number of new non-games). Changes in the revenue streams of paid and non-paid apps (e.g., the increasing prevalence of in-app purchases) may result in changes in the number of entrants into the market. Such changes could also drive app entry and undermine the search mechanism explanation. This does not appear to be the case in the data. Panel (c) shows that there are no changes in the patterns of free and paid product entry between games and non-games after re-categorization.[20]

In addition to the difference-in-differences estimates, I also test for whether changes in the number of other apps in a category affect a paid app's prices. As in Column (3) of Table 3 I use short run changes in the number of apps in game

[19] With lower costs, higher valuation consumers can discover more preferred game-apps more easily (Bar-Isaac et al. 2012).

[20] There are substantial changes in the absolute share of paid apps that are entering into the market over time. The share of new paid products falls from over 30% to less than 10%. This pattern is consistent for both games and for non-game apps.

Table C10: Prices: Difference-in-Differences Estimates

Panel (a): All Paid Apps

Outcome Variable:	Mean Price	Mean Price	Mean Price	Mean Price
Games × Post	0.037	-0.037	-0.147	0.019
	(0.036)	(0.122)	(0.152)	(0.105)
Games × Post × No Pre-Existing			0.197	
			(0.138)	
Games × Post × Small Type				-0.595
				(0.347)
Games	9.966			
	(0.758)			
Observations	70	1,470	1,470	980
R-squared	0.999	0.970	0.970	0.966

Panel (b): Paid Entrant Apps

Outcome Variable:	Mean Price	Mean Price	Mean Price	Mean Price
Games × Post	0.379	0.076	0.030	0.000
	(0.443)	(0.423)	(0.470)	(0.459)
Games × Post × No Pre-Existing			0.082	
			(0.288)	
Games × Post × Small Type				-0.050
				(0.910)
Games	28.088			
	(10.409)			
Observations	70	1,470	1,470	980
R-squared	0.999	0.970	0.970	0.966
Unit of Observation:	Game/Non-Game	App-Type	App-Type	App-Type
Sample:	All Paid	All Paid	All Paid	All Paid
Sample Period:	Jan 12/Dec 14	Jan 12/Dec 14	Jan 12/Dec 14	Jan 12/Dec 14
Year/Month FE	•	•	•	•
App-Type FE		•	•	•

Notes: Sample period in all columns covers January 2012-December 2014. Sample in Column (1) consists of monthly observations at the Game/Non-Game level. Sample in Columns (2)-(4) consists of monthly observations at the app-type level. Outcomes in panel (a) are average prices for all paid apps at each aggregation level. Outcomes in panel (b) are average prices for all paid new entrants at each aggregation level. Controls include year and month fixed effects, and game/non-game fixed effects or app-type fixed effects, depending on the column. Additional controls include game/non-game or app-type time trends. The variable "Games × Post" is a dummy variable equal to 1 for games, or game app-types during and after March 2014. Standard errors are robust to heteroskedasticity in Column (1) and are clustered at the app-type level in the remaining columns.

categories due to re-categorization. Apps move from being in large categories with many other apps of different types to smaller categories with other apps of their own type. One explanation for the findings in the main text showing that downloads for games with more apps in their category fall is that these apps face more competition from imperfect substitutes. If this is the case, there should also be a link between the number of apps in a category and prices.

In Table C11 I show the results of a regression relating pre/post re-categorization differences in the number of apps in game categories to pre/post differences in individual app prices. Coefficient estimates are both small in absolute terms and are statistically null. There is no relationship between changes in the number of apps in a category and changes in app prices. This suggests that changes in the number of apps in a category does not affect competition. Instead, it primarily affects the market by reducing congestion.

Table C11: Price Changes in Response to Changes in Number of Apps in a Category

	(1) Post/Pre Δ ln(Price)	(2) Post/Pre Δ Price
Post/Pre Δ ln(N Apps in Category)	-0.000	-0.005
	(0.002)	(0.009)
Observations	21,749	21,749
R-squared	0.449	0.272
Unit of Observation:	App	App
Sample Period:	Jan 14 / Apr 14	Jan 14 / Apr 14
Sample:	All Paid Games	All Paid Games
App Controls	•	•

Notes: Sample period in both columns covers January 2014 to April 2014. Sample includes monthly observations of all paid game apps present from January 2014 to April 2014. Additional app-level controls include average app ratings and app age-specific fixed effects. The outcomes are differences between app-level average price in March and April 2014 and app-level average price in January and February 2014. Δ ln(N Apps in Category) is the difference in the natural log of the number of apps in the category of app j after re-categorization (March and April 2014) and the natural log of the number of apps in the category of app j before re-categorization (January and February 2014). Standard errors are robust to heteroskedasticity.

C.9 Google Trends Evidence of Consumer Awareness of Android Games and Non-Games

I do not observe Google's advertising for the Google Play app store. Instead, I use Google Trends search volumes to proxy consumer awareness for Android Games and Android Apps (Google 2014). Figure C5 shows the weekly Google Trends volumes from January 2012 to December 2014. The top two panels compare Google Trends for the "Android Games" and "Android Apps" search queries. The middle panels compare "Google Play Games" and "Google Play Apps." The last two panels compare "Google Play Games" and "iOS Games." In all cases Google trends are measured relative to the maximum search volume over the period. The figures on the left are absolute search trends numbers and the figures on the right are search trend ratios. Google Play/Android Games volumes are always the numerators in the ratios.

The figures all show that there is substantial variation in search query volumes over the sample period. For example, there is a spike in search queries around

Christmas/the New Year. In all three sets of comparisons there is no spike in Google Play/Android search queries around the period of the re-categorization of the store (solid vertical red line). There is also no change in the relative search query ratio around the period of the re-categorization. In the first two panels (Android Games vs. Android Apps), the "Google Play Games" search query is trending upwards relative to the "Google Play Apps" search query. There is no change in this trend around the re-categorization period.

C.10 Developer Switching Between Games and Non-Games

Figure C6 shows a ratio of the number of existing non-game developers[21] who produce a game in period t over the number of existing game developers who produce a non-game in period t. The ratio is greater than 1: there are more non-game developers switching to producing games than the other way around. The figure also shows that there is an increase in the ratio around the period of the game re-categorization (from 2 to 2.6). This is potentially consistent with a resource allocation story whereby developers have a fixed budget and have to choose between producing games and non-games. However, within 2 months of the re-categorization, the ratio falls to pre re-categorization levels.

This suggests that developer switching after re-categorization is a short term response. By comparison, the entry effects of re-categorization are a long term phenomenon. Period-specific treatment effects captured in Figure C3 show that the increase in the number of games relative to non-games persists all the way to the end of the sample (9 months after re-categorization). The magnitude of the treatment coefficient in the last month of the sample in Figure C3 is as large as the magnitude in the second month after re-categorization. This suggests that the magnitude of the treatment effect cannot be explained by developers switching from producing non-games to producing games.

D Appendix D

D.1 Additional Demand Model Parameter Estimates

This section shows and discusses additional coefficient estimates of demand estimates from Column (4) of Table 4. Table D1 shows estimates for coefficients on lagged

[21] Defined as those who only developed non-games in the past.

downloads (q_{jt-1}), and various proxies for app quality - the number of screenshots, app size in MB, and whether or not an app has a video preview.

Table D1: **Additional Table 4 Column (4) Parameter Estimates**

γ Estimates	
ln(Lag App Downloads)	0.034
	(0.004)
ln(Size in MB)	0.046
	(0.006)
N Screenshots	0.008
	(0.001)
Video Preview Dummy	0.072
	(0.017)
Paid App Dummy	0.107
	(0.194)

The coefficient on lagged downloads is positive. It suggests that apps with more past downloads are easier to find by consumers. It is also consistent with previous findings in the empirical literature on online product ranking-based discovery frictions (e.g., Ursu 2018). The positive coefficients on variables capturing app quality generally go in the expected direction. more screenshots, a video preview, and bigger app size should reflect higher app quality and generate additional consumer utility.

Figure D1 plots estimates of app-type specific differences in pre/post re-categorization fixed effects. It shows substantial heterogeneity across app-types. On average, the ten app-types that did not have a pre-existing category (Adventure, Board, Educational, Family, Music, Role Playing, Simulation, Strategy, Trivia and Word) experience larger average increases in utility as compared to the eight app-types that had explicit categories before (Action, Arcade, Card, Casino, Casual, Puzzle, Racing and Sports).[22] These effects are quantitatively large: on average, utility increases by 1 dollar for consumers from buying an app belonging to an app-type that did not have a category before the change, holding everything else constant.[23] This is consistent with reduced form evidence from Section C.2, showing that re-categorization

[22]The main exception for this group is Action games, which have a change in fixed effects comparable to some of the other app-types. This is possibly because it was grouped together with the Arcade app-type before re-categorization, leading to substantial improvement in the consumer search process.

[23]Re-categorization also increases average utility for consumers from purchasing other app-types, although effects there are less than half the size on average (except for Action games, see previous footnote). This is likely because the informativeness changed for the other app-types as well, albeit

increased downloads for those app-types more than for the second group of app-types, and suggests that informativeness of the category structure increased after re-categorization.

In Appendix D.4 I show that this change is driven by the re-categorization, rather than some other average app-type time-varying differences.

D.2 Demand Model with Search

This section describes a demand model with search following the consideration set approach of Moraga-González et al. (2015).[24] Consumers choose a single product out of a set of N products. For each product j, consumers obtain utility $u_{ij} = \delta_j + \epsilon_{ij}$. Consumers are not fully informed about products: they do not know their ϵ_{ij}s. Search resolves this uncertainty. Consumers in this market first choose a consideration set A of products and pay a set-specific search cost. They find out the ϵs of those products and pick a product j out of subset A. In this application, products in subset A can be located across multiple categories. Subsets are unobserved to the econometrician. Consumers know the expected utility (or inclusive value) they obtain from the products in subset A: $\overline{U}_A$.[25] Consumers incur subset-specific search costs (c_{iA}) such that the utility of consumer i choosing subset A is:

$$u_{iA} = \overline{U}_A - c_{iA} = \overline{U}_A - \left(\sum_{r \in A} \theta\psi_r + \lambda_{iA}\right) \tag{4}$$

where ψ_r reflects a deterministic "distance" between the consumer and product r in set A. λ_{iA} is a consumer/choice set specific search cost shock, which I assume is EV type 1 distributed mean zero with a standard error normalized to 1.[26] This shock can be interpreted as an information shock - word of mouth from friends or family. θ is effectively the average marginal search cost for consumers in the market. As with consumer utility, search costs have no unobservable heterogeneity aside from the idiosyncratic shock.

Due to the idiosyncratic error terms on both search costs and consumer utility, the unconditional probability of a consumer choosing product j is:

in relatively minor ways. For example, a consumer looking for a Card game may be less confused about what kinds of Card games are in the "Cards" category (i.e., no family card games, no music card games).

[24]It is also similar to Goeree (2008) and Honka et al. (2017).

[25]In a multinomial logit model, this is simply $log[1 + \sum_{r \in A} exp(\delta_r)]$. Consumers also always have the outside option, regardless of the set they consider.

[26]θ can vary across products or product groups.

$$P_j = \sum_{A \in A_j} P_A P_{j|A} \tag{5}$$

where A_j is the set of all subsets that product j belongs to, P_A is the probability of a consumer choosing subset A From the set of all possible subsets and $P_{j|A}$ is the probability that consumer i picks product j out of subset A. The unconditional probability P_j is equivalent to the observed market share of product j (s_j). Moraga-González et al. (2015) show that it is possible to "integrate out" the unobservable subsets and obtain the following closed form expression for s_j:[27]

$$s_j = \frac{\frac{exp(\delta_j)}{1+exp(\theta\psi_j)}}{1 + \sum_{k \in N} \frac{exp(\delta_k)}{1+exp(\theta\psi_k)}} \tag{6}$$

where the denominator sums up over all products in the market (N) rather than over specific subsets. This expression is effectively the standard multinomial logit model except that the market share of product j is shaded down by how hard it is to find (ψ_j). I include the "discovery cost" variables from the model in the main text ($N_{c^*(j)}$ and R_{jc}) in ψ_j. The resulting expression is similar to the market share specification in Equation 4 in the main text. Setting $exp(\gamma ln(N_{c^*(j)}) + R_{jc}\kappa) = \frac{1}{1+exp(\theta\psi_j)}$ and introducing an additional nested logit error term equates the two.[28]

I estimate this model using non-linear GMM with the same instruments used to estimate the linear demand model in the main text. Parameter estimates from this model are in Column (3) of Table D3. These are qualitatively similar to demand estimates from the main text. Note that signs for the "search cost" parameters are flipped relative to results in the main text because of how they enter into the model.

While this is a reasonable approach to modelling consumer product discovery and demand in the mobile app market, there are potentially many other ways in which consumers search the market. This model also does not easily allow controlling for additional unobservable heterogenity with aggregate product level data. I choose to use the simpler linear demand model in the main text. It does not make specific assumptions about the consumer search process, but is broadly consistent with many predictions from search literature.

[27] The assumption that the "distance" of products in a consideration set is additive in the set's search costs is key for obtaining a closed form expression for choice probabilities.

[28] This consideration-set based model allows for unobservable heterogeneity in consumer preferences, such as a consumer/category specific shock. However, the standard market-share inversion procedure for nested logit models does not apply to the consideration set model, and it would have to be estimated by simulation. I do not include additional unobservable heterogeneity for this reason.

D.3 Additional Demand Model Regressions

Table D2: **1st Stage Supporting Regression**

	(1)
Outcome Variable:	ln(N Category Apps_t)
ln(N Category Apps_{t-1})	0.142
	(0.022)
Mean Category Rating_t	-0.279
	(0.065)
ln(Category Downloads_t)	0.387
	(0.034)
Category FE	•
Observations	624
R-squared	0.953

Notes: The sample includes monthly category-level observations from February 2012 to December 2014. Standard errors are robust to heteroskedasticity.

D.4 Placebo Time-Varying Fixed Effects

In the main text, I include two sets of time-varying fixed app-type fixed effects: a set of app-type fixed effects that turns on before re-categorization takes place, and a set of app-type fixed effects that turns on after re-categorization takes place. The difference in these fixed effects is in Figure D1 in Appendix D.1. It shows app-type level welfare changes. I interpret these changes as being driven by re-categorization improving information, but they could also be caused by other time varying heterogeneity within app-types. For example, Educational games have the biggest fixed effect change, which could be the result of consumers liking educational games more over time.

To test whether the variation in fixed effects is driven by the re-categorization event, I introduce a specification of the model with three sets of time-varying app-type fixed effects. The first set of app-type fixed effects is active from March 2012 to February 2014. The second set of app-type fixed effects is active only during April 2014 (March 2014 is omitted from the data) and the last set of app-type fixed effects is active from May 2014 to December 2014. The change between the first two sets identifies information effects just around re-categorization. The change between the second two sets identifies whether there were other changes over time. If changes in app-type fixed effects primarily capture changes in category informativeness, I should

Table D3: **Additional Demand Estimates**

	(1)	(2)	(3)
γ		-0.392	0.374
		(0.026)	(0.016)
$\gamma\times$ New App		-0.012	
		(0.011)	
σ	1.404	0.708	
	(0.069)	(0.027)	
β_{price}	-3.800	-0.833	-1.470
	(0.119)	(0.111)	(0.041)
ln(Lag App Downloads)		0.036	-0.282
		(0.004)	(0.008)
ln(Size in MB)	0.078	0.046	0.099
	(0.007)	(0.006)	(0.002)
Video Preview Dummy	-0.016	0.072	0.138
	(0.020)	(0.017)	(0.005)
N Screenshots	0.000	0.008	0.022
	(0.002)	(0.001)	(0.000)
Paid and New App Dummies	•	•	•
App Age FE	•	•	•
App Rating FE	•	•	•
Year/Month FE	•	•	•
Developer FE	•	•	•
App Type $\times$ Pre/Post Re-Categorization FE	•	•	•
Observations	4,152,147	4,152,147	4,167,060

Notes: The sample includes monthly observations of all free and paid game apps in the Google Play Store from March 2012 to December 2014, excluding March 2014. Column (1) shows estimates of a nested logit model without discovery friction controls. Column (2) shows estimates of the model from the main text with heterogeneous discovery frictions for new and incumbent apps. Column (3) shows estimates of the search and demand model described in Online Appendix D.2. "App Rating FE" are a set of dummies representing the average rating of app j in period t within 0.5 stars. Apps with 2 stars or less are the "baseline" category for "App Rating FE." $Year/MonthFE$ include year and month dummies. Instruments for price and for σ include the ratings of other apps in the same category, the number of screenshots of other apps of the same app-type and the average size of other apps of the same app-type. Instruments for lagged downloads for app j include functions of further lags in app j downloads (2 and 3 periods before period t). The instrument for the number of apps in the category is described in the main text and is the residual of the regression in Table D2. Standard errors are clustered at the app level in Columns (1) and are (2) and robust to heteroskedasticity in Column (3).

not see substantial differences between the April 2014 and May-December 2014 fixed effects where informativeness was constant.

Figure D2 shows both sets of differences for each app-type and their computed 95% confidence intervals. The results suggest that consumer utility from app types changed around the period of re-categorization and not after. Most of the "placebo" fixed effect differences are statistically zero at the 95% confidence level. Even for the few app-types where these differences are not statistically zero, they are very small in magnitude relative to the true difference in fixed effects between the pre- and post-re-categorization.

D.5 Distributions of Main Model Welfare Effects

Figure D3 shows the full distribution of welfare outcomes generated by the randomized simulations. The figures show that the randomizations do not substantially change the main effects.

Online Appendix References

Bar-Isaac, H., G. Caruana, and V. Cuñat Martinez (2012). Search, design and market structure. *American Economic Review 102*(2), 1140–1160.

Chevalier, J. A. and D. Mayzlin (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research 43*(3), 345–354.

Eeckhout, J. (2004). Gibrat's law for (all) cities. *American Economic Review 94*(5), 1429–1451.

Garg, R. and R. Telang (2013). Inferring app demand from publicly available data. *MIS Quarterly 37*(4), 1253–1264.

Goeree, M. S. (2008). Limited information and advertising in the us personal computer industry. *Econometrica 76*(5), 1017–1074.

Google (2012-2014). Google trends of searches for android and ios games and apps. Downloaded from "https://www.google.com/trends/". Accessed in 2019.

Honka, E., A. Hortaçsu, and M. A. Vitorino (2017). Advertising, consumer awareness, and choice: Evidence from the us banking industry. *The RAND Journal of Economics 48*(3), 611–646.

Kummer, M. and P. Schulte (2019). When private information settles the bill: Money and privacy in google's market for smartphone applications. *Management Science 65*(8), 3470–3494.

Leyden, B. T. (2018). There's an app (update) for that. mimeo.

Liebowitz, S. J. and A. Zentner (2020). The challenges of using ranks to estimate sales. *Available at SSRN 3543827*.

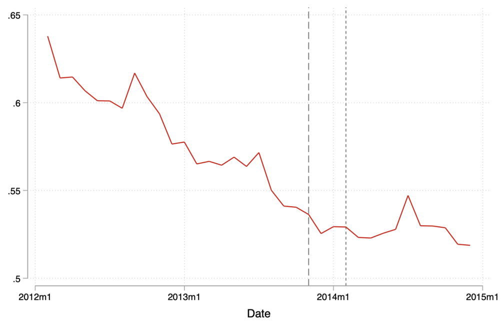
Liu, Y., D. Nekipelov, and M. Park (2014). Timely versus quality innovation: The case of mobile applications on itunes and google play. *NBER Working Paper*.

Moraga-González, J. L., Z. Sándor, and M. R. Wildenbeest (2015). Consumer search and prices in the automobile market.

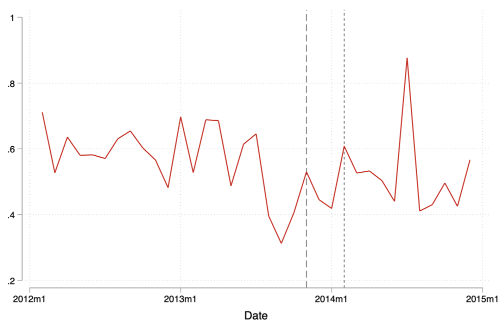
Ursu, R. M. (2018). The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions. *Marketing Science 37*(4), 530–552.

Figure C4: Prices

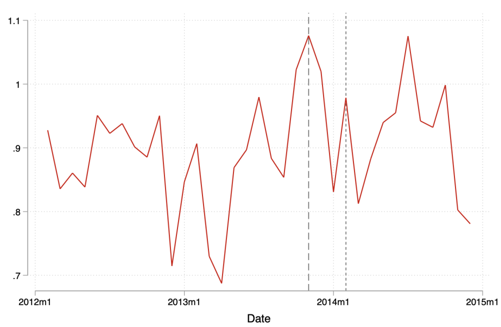
(a) Outcome: Mean Game Price / Mean Non-Game Price



(b) Outcome: Mean Entrant Game Price / Mean Entrant Non Game Price



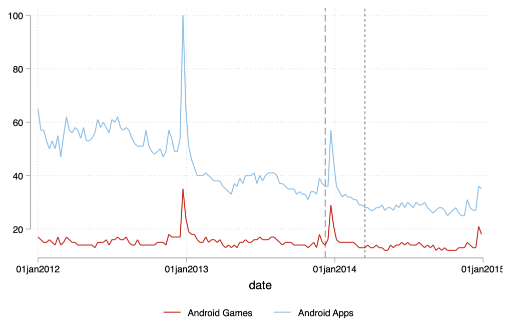
(c) Outcome: Share Paid Game Entrants / Share Paid Non-Game Entrants



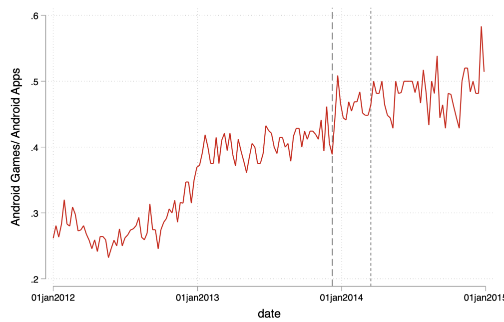
Notes: Panel (a) shows a ratio of mean monthly game app price over mean monthly non-game app price using all paid apps. Panel (b) shows a ratio of mean monthly game app price over mean monthly non-game apps price using only paid entrants. Panel (c) shows a ratio of the monthly percentage of new game apps that are paid over the monthly percentage of new non-game apps that are paid. In all panels, the first dashed vertical line represents the re-categorization announcement and the second dashed vertical line represents the start of the re-categorization period.

Figure C5: US Google Search Trends

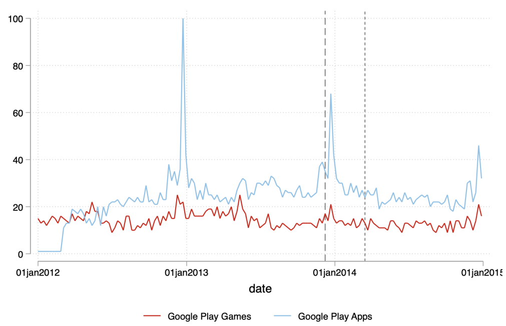
(a) "Android" Searches



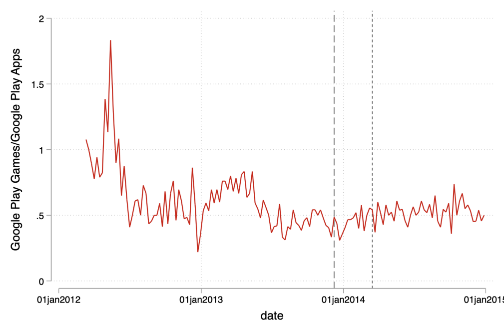
(b) Ratio of Panel (a) Games/Apps



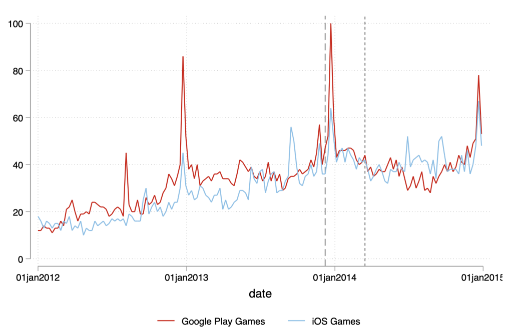
(c) "Google Play" Searches



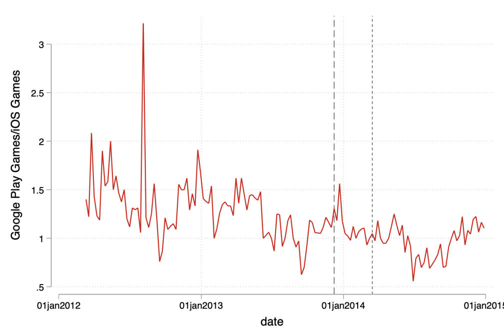
(d) Ratio of Panel (c) Games/Apps



(e) Google Play/iOS Searches

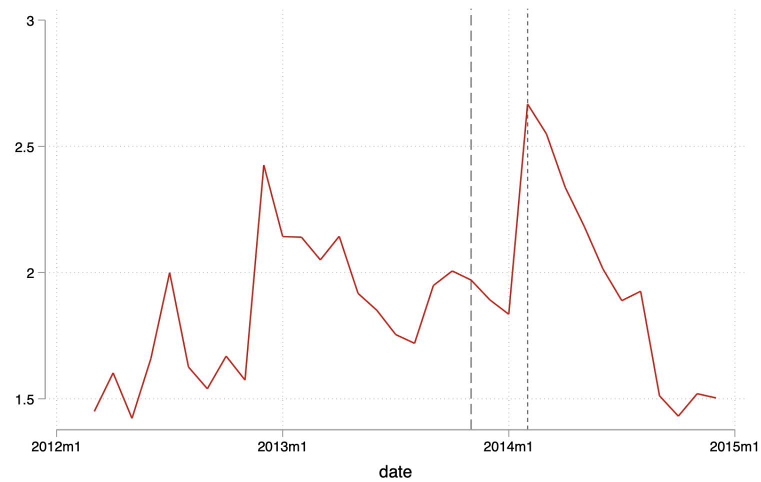


(f) Ratio of Panel (e) Google Play/iOS



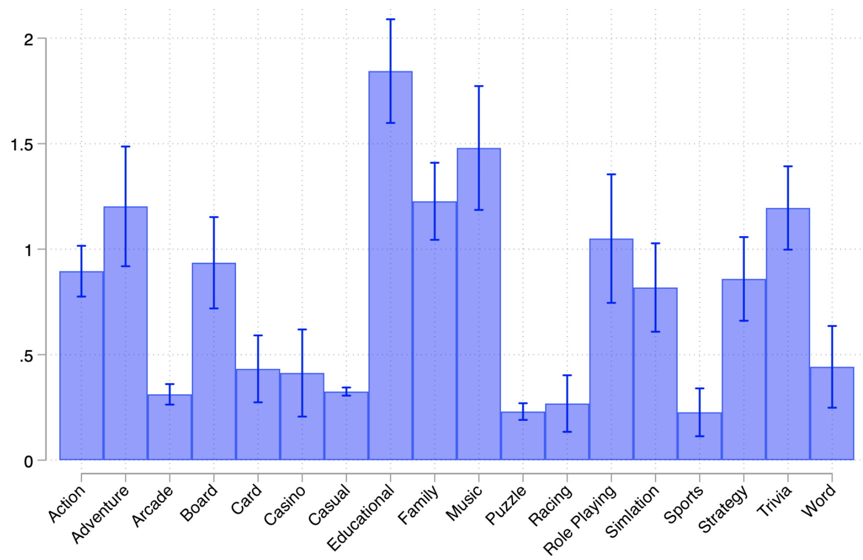
Notes: Panels (a), (c) and (e) show daily Google Trend search volume estimates for different queries. In each of the panels, numbers are normalized relative to maximum search volume which is set to 100. Panels (b), (d) and (f) show ratios of the numbers in panels (a), (c) and (e), respectively. In all panels, the first dashed vertical line represents the re-categorization announcement and the second dashed vertical line represents the start of the re-categorization period.

Figure C6: Ratio of Switching Developers: $\frac{\text{Non-Game to Game}}{\text{Game to Non-Game}}$



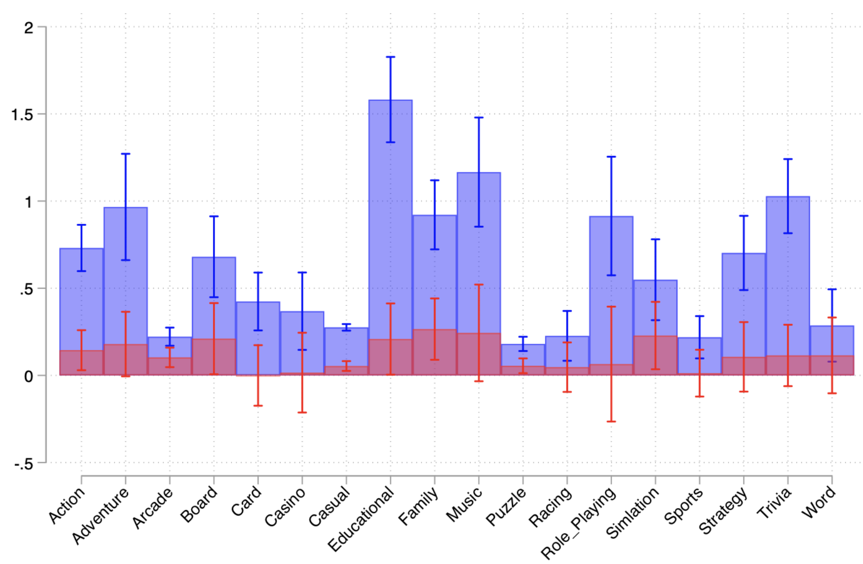
Notes: This figure shows a monthly ratio. In each month t the numerator is the number of developers who produced a non-game app in any period before t and produced a game app in period t. The denominator is the number of developers who produced a game app in any period before t and produced a non-game app in period t. The first dashed vertical line represents the re-categorization announcement and the second dashed vertical line represents the start of the re-categorization period

Figure D1: Difference in App-Type Fixed Effects



Notes: Each column shows the difference in estimated app-type fixed effects based on the model described in Section IV: the fixed effect for app-type c in the post- re-categorization period, minus the fixed effect value for app-type c in the pre- re-categorization period. 95% calculated confidence interval for this difference is shown.

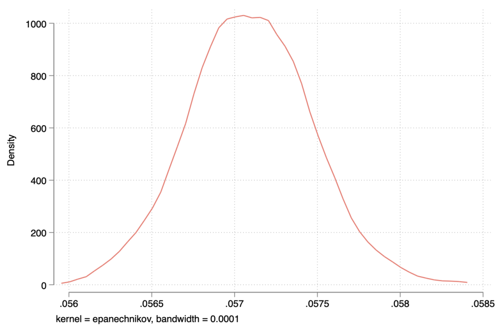
Figure D2: Differences in App-Type Fixed Effects



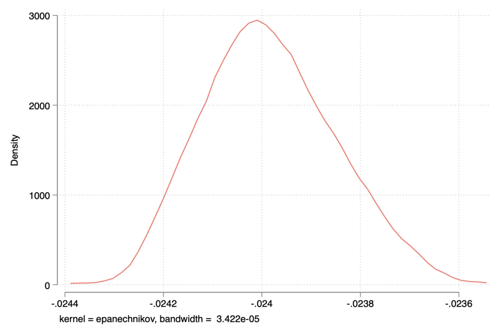
Notes: Each column shows two sets of differences in estimated app-type fixed effects based on the model described in Section IV. The first difference, in blue, is between the app-type fixed effect for Mar 2012-Feb 2014 and the app-type fixed effect for April 2014. The second difference, in red, is between the app-type fixed effect for April 2014 and the app-type fixed effect for May 2014-December 2014. 95% calculated confidence intervals for both differences are shown.

Figure D3: Distribution of Welfare Effects Across Simulations

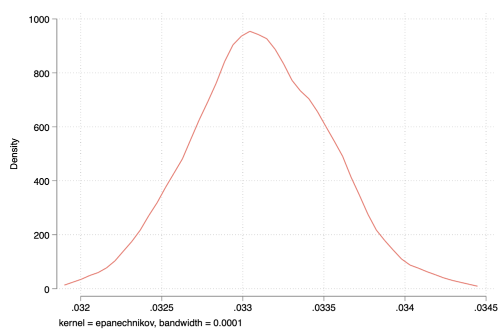
(a) Gross Welfare Change
(mean = 0.057)



(b) Congestion Cost Change
(mean = -0.024)



(c) Net Welfare Change
(mean = 0.033)



Notes: Panels show the distribution of welfare effects across 500 simulations.