Internet Appendix for "Gatekeepers and Kingmakers: Do Superplatforms Distort Digital Competition?"

Umit G. Gurun, Zheng Liu, Steven Chong Xiao

A Additional Empirical Results

A.1 Supporting Figures and Tables

This section provides additional figures and tables referenced in the main paper.

Distribution of negative review removal rates. Figure A.1 illustrates the distribution of negative review removal rates for GV-backed and other-VC-backed apps across Android and iOS platforms. The horizontal axis represents the removal rate, calculated as the proportion of negative reviews removed relative to the total number of initially collected negative reviews. For Android apps (subfigure (a)), the distribution for GV apps is noticeably more left-skewed than for other-VC apps, indicating that a higher proportion of GV apps exhibit elevated removal rates for negative reviews. In contrast, for iOS apps (subfigure (b)), GV apps show a greater concentration around a 0% removal rate relative to their Android counterparts. Complementary evidence in Table A.1 reveals that 13.16% of GV Android apps have at least 10% of their negative reviews removed, significantly higher than the 3.69% rate for other-VC Android apps (difference significant at the 5% level). No statistically significant differences in the negative review removal rates emerge between GV apps and other-VC apps on Apple's App Store. These findings suggest that GV-backed apps are associated with more aggressive curation of negative reviews on Google Play, which is not evident on the other platform.

Content analysis of removed reviews. One potential alternative explanation for the observed review curation pattern is that GV apps may feature a higher proportion of inappropriate reviews that violate Google Play's policies, thereby inducing more platform removals. To investigate this, we train a neural network classifier on removed and retained negative reviews from other-VC apps and then apply it to GV app reviews. Table A.2 presents the model performance. On the training and validation sets of other-VC app reviews, the model achieves consistent accuracy (61% and 61%), precision (62% and 62%), and recall (57% and 57%), indicating stable criteria for review removal across other-VC apps. However, when applied to GV app reviews, performance drops to 53% accuracy, 55% precision, and 36% recall, suggesting that the criteria used to remove negative reviews from GV apps differ from those applied to other-VC apps.

Ranking of VC treatment effects. Figure A.2 provides a visual representation of the percentile rankings of treatment effects for each VC across three app performance metrics. GV ranks in the 91st percentile for review volume treatment effects, the 90th percentile for rating treatment effects, and the 98th percentile for bug-reduction treatment effects. In terms of the average percentile across these three dimensions, GV ranks as the best-performing VC among all 128 VCs in the sample. This evidence supports our identification assumption and our causal interpretation of the DiD estimates presented in the main paper.

App update frequency and quality improvements. An alternative explanation for the observed GV treatment effects is that they stem from technical guidance provided by GV that enhances the quality of Android apps. Given Google's role in designing the Android operating system, it is reasonable to presume that GV possesses specialized expertise in Android app development. If GV-backed Android apps do indeed undergo quality improvements, these enhancements should manifest through periodic app updates, which are crucial for addressing bugs, bolstering security, and introducing new features. Table A.3 estimates Model (3) using update frequency as the dependent variable. The coefficient on *Post* $\times$ *Android app* is statistically insignificant across all specifications, indicating that GV-backed

startups do not update their Android apps more frequently compared to their iOS counterparts after receiving GV investment.

We further examine whether major app updates drive the GV treatment effects. We replace the continuous *Update Frequency* measure in the triple-interaction specification of Table 8 with a binary indicator, *Major Update*, which captures whether an app underwent a primary version change (e.g., from version 1.7.8 to version 2.0.1). Such primary version changes typically reflect more substantial updates beyond minor fixes. As shown in Table A.4, the triple interaction term between *Post*, *Android app*, and *Major Update* remains mostly insignificant. Combined with the results in Table 8 in the main paper, these findings suggest that intrinsic quality improvements through technical upgrades are unlikely to drive the observed performance gains for GV-backed apps.

A.2 GV's Role in the Funding Syndicate

We examine whether GV's self-preferencing behavior is reflected in its syndication patterns and whether the treatment effects differ based on GV's role as lead versus co-investor.

Co-investment patterns. If the Google Play platform provides a form of risk mitigation for GV's portfolio companies through preferential treatment, GV may have less need to share risk through larger investment syndicates. We investigate this by comparing co-investment patterns for GV-backed startups with and without mobile apps in our sample. Table A.5 presents the results based on all GV-invested financing rounds from 2008 to 2022. The sample includes 692 financing rounds for GV-backed startups without matched apps and 146 financing rounds for GV-backed startups whose apps appear in our sample. GV-backed startups with apps have, on average, 4.81 co-investors per round, compared to 5.36 for GV-backed startups without matched apps. This difference of 0.55 co-investors is statistically significant at the 5% level (T = 2.29). The effect is modest in magnitude, which is reasonable since co-investment decisions are driven by multiple factors beyond platform support, including deal size, sector expertise, and the startup's financing needs. Nevertheless, this finding is consistent with the notion that platform-mediated support may partially substitute for the risk-sharing function typically performed by a larger syndicate of co-investors.

Lead versus co-investor status. A common distinction in the VC literature is between lead investors, who typically take board seats and provide hands-on guidance, and follow-on investors, who may be more passive. Our treatment variable combines rounds in which GV served as the lead investor with those in which it participated as a co-investor. This is appropriate in our setting because the hypothesized mechanism, i.e., platform-mediated self-preferencing, operates through Google's control over Google Play's recommendation algorithms and content moderation policies, which should be available regardless of whether GV led the financing round or participated alongside other investors.

Empirically, we test whether the treatment effect differs by GV's lead-investor status. Among the 17 startups in our DiD sample, GV served as the lead investor in the first round it participated in for 9 startups, while it participated as a co-investor for the remaining 8. Table A.6 augments our baseline DiD model with a triple interaction between Post, Android app, and a Lead VC indicator. The triple interaction terms are not statistically significant for any of the three performance metrics: 0.677 (T = 1.33) for log review volume, −0.107 (T = −0.38) for mean rating, and −0.005 (T = −0.21) for %Bug. This null result is consistent with our interpretation that the treatment effects operate through platform-level interventions rather than investor-specific expertise, and supports our choice to combine lead and follow-on investments in the baseline analysis.

A.3 Self-Preferencing Across Investment Stages and Exit Status

We examine whether GV's self-preferencing varies across investment stages and between current and exited portfolio companies.

Investment stage analysis. We classify financing stages based on each startup's most recent financing round before the end of 2022 in the LSEG Workspace database. Among the 38 GV apps in our review curation sample, 7 are developed by early-stage startups, 14 by expansion-stage startups, and 17 by later-stage startups. Table A.7 compares review curation across these groups. The directional pattern suggests more pronounced curation for early-stage GV apps: early-stage apps experience a 7.42% decline in negative reviews (vs. 4.92% for later-stage apps) and a 12.43% decline in bug-related reviews (vs. 2.88%). However, the differences are not statistically significant ($T = -0.61$ for negative reviews; $T = -1.49$ for bug-related reviews), reflecting the limited statistical power after further subsetting the sample by financing stage.

Table A.8 augments the cross-sectional visibility regressions with a GV $\times$ Early stage interaction. The interaction coefficients are positive across all four visibility measures (ranging from 0.861 to 1.461), with statistically significant effects for overall and cross-category visibility. In economic terms, the estimates imply that early-stage GV apps receive an additional 331% overall visibility advantage (i.e., $e^{1.461} - 1$) and an additional 231% cross-category visibility advantage (i.e., $e^{1.196} - 1$).

We also estimate a triple-difference model within our DiD framework. The early-stage dummy equals one for app-quarter observations before the startup's most recent early-stage financing round. Of the 17 startups in the DiD sample, 8 received GV investment while still in the early-stage phase. As shown in Table A.9, the baseline treatment effects (Post $\times$ Android app) remain robust. The triple interaction terms (Post $\times$ Android app $\times$ Early stage) are not statistically significant, as expected given the limited variation with only 8 early-stage startups among 17.

Current versus exited portfolio companies. To test whether self-preferencing serves as a precursor to acquisition or other exits, we compare review curation for apps of currently invested GV startups (38 apps) against apps of startups that had already exited GV's portfolio by the end of 2022 (15 apps). Because this end-of-2022 exit cutoff predates the July 2023 no-exit screen that defines the currently invested group, startups exiting between January and July 2023 enter neither group. The 15 exited apps are not among the 38 GV apps in our review curation sample, which includes only apps of startups still in GV's portfolio during the data collection period; they enter this analysis as a separate comparison group. The measures are the same cross-snapshot review-removal rates used in Table 4: for each app, we compute the percentage change between the two Google Play snapshots in the volume of reviews posted from July 2022 to June 2023, separately by review category, and compare group means for current and exited GV apps. Table A.10 shows that current GV apps exhibit slightly more curation (e.g., -5.21% vs. -3.68% for negative reviews; -4.71% vs. -3.12% for bug-related reviews), but the differences are not statistically significant ($T = -0.80$ and $T = -0.87$, respectively). The lack of statistical precision has two natural sources. First, the exited group contains only 15 apps, so the group-difference tests have limited power. Second, exit status is measured as of the end of 2022, close to the July 2022 to June 2023 review period over which removal is observed, so for recently exited startups there is little time for the platform's treatment of their apps to have adjusted.

Interpreting the comparison also requires care about what "exit" means in our data. Exit status is drawn from LSEG Workspace and aggregates several event types, including mergers, IPOs, secondary sales, and other exit events (see Section A.6); we do not split the exited group by exit type because 15 apps cannot support such subsamples. The distinction matters for the directional prediction. After an acquisition, GV's stake in the startup is extinguished, and the precursor-to-exit logic predicts

that curation should fade. After an IPO, by contrast, GV may continue to hold shares for some time, and while it does, it retains a financial incentive to support the app's perceived quality, so self-preferencing need not stop at the exit date. This distinction helps explain why the post-exit estimates, while smaller than those for current portfolio companies, remain economically meaningful rather than reverting to the other-VC benchmark: that pattern is consistent with post-IPO holding incentives, with gradual unwinding of preferential treatment after exit, or with persistence in how the platform's recommendation and moderation systems treat previously favored apps.

Overall, while the inherent power limitations of our setting prevent us from establishing these distinctions statistically, the directional consistency across all subsplits — early-stage apps showing more curation than later-stage apps, and current portfolio companies showing more curation than exited ones — points in the predicted direction, lending qualitative support to the conjecture that self-preferencing intensity varies with the platform's financial stake and strategic interest in the startup.

A.4 Robustness to Staggered Treatment Timing

Our DiD analysis spans a long sample period (2010–2023) with GV investments occurring at different times across startups. While our empirical design, which compares Android and iOS versions of the same app within app-pair × calendar-quarter fixed effects, is essentially a stacked DiD that mitigates many of the concerns raised in the modern DiD literature (Callaway and Sant'Anna, 2021; Sun and Abraham, 2021; Wing et al., 2024), we conduct two additional robustness tests to further address potential issues associated with differential treatment timing.

Event window restriction. We restrict the sample to a symmetric window of 12 quarters (3 years) before and after each startup's GV investment. This restriction reduces exposure to slow-moving secular changes over the 2010–2023 period and limits identification to comparisons close to the event. As shown in Table A.11, the results are robust to this restriction. The coefficient on Post × Android app remains positive and significant for log review volume (0.325, T = 1.76) and mean rating (0.284, T = 2.42), and negative and significant for %Bug (−0.031, T = −2.06). These results confirm that our baseline findings are not driven by long-run secular trends or period-specific confounds.

Two-stage difference-in-differences. We implement the two-stage DiD estimator proposed by Gardner (2022), which estimates fixed effects in a first stage using only untreated observations and then regresses outcomes net of these fixed effects on treatment status in a second stage. Under parallel trends, this procedure identifies an interpretable overall average treatment effect on the treated even with staggered timing and heterogeneous effects. In our setting, the iOS versions of each app provide a natural untreated comparison series throughout the sample period, supporting the first-stage estimation. Table A.12 shows that the review-volume result is robust: the coefficient on Post × Android app is 0.920 (Z = 5.22, $p < 0.01$), consistent with and slightly larger than the baseline estimate. For mean rating and the proportion of bug-related reviews, the point estimates are signed consistently with the baseline (0.077 and −0.008, respectively) but do not reach statistical significance.

Overall, these analyses confirm that our baseline findings are robust to concerns about the long sample horizon and staggered treatment timing.

A.5 Startup Financing Outcomes

Table A.13 presents the financing outcomes for startups affiliated with GV and other VCs during the sample period. Overall, GV-invested startups exhibit more favorable financing outcomes. Specifically, GV-invested startups participate in an average of 5.95 funding rounds, which is 3.28 rounds more than

those financed by other VCs, a difference statistically significant at the 1% level. These startups also secure a cumulative average financing of $639.46 million, in contrast to the $30.29 million accrued by startups backed by other VCs, with this difference also significant at the 5% level. Moreover, on a per-round basis, GV-invested startups receive an average of $72.86 million, compared to $11.16 million for startups associated with other VCs, a difference significant at the 1% level.

Has the superior performance of GV apps contributed to better financing outcomes for GV-invested startups? To shed light on this question, we construct a startup-year panel dataset, pairing each startup with its most downloaded app on Google Play. Then, we estimate the following regression to examine the correlation between app performance and firm financing outcomes:

$$\begin{aligned}\text{Financing outcome}_{i,t+1} =& \beta_0 + \beta_1 \Delta\text{Log(review volume)}_{i,t} + \beta_2 \Delta\text{Mean rating}_{i,t} \\ &+ \beta_3 \Delta\%\text{Bug}_{i,t} + \epsilon_{i,t}\end{aligned} \tag{A.1}$$

The dependent variable, *Financing outcome*$_{i,t+1}$, denotes various measures of financing outcome for startup i in year $t+1$. Δ*Log(review volume)*$_{i,t}$ is the change in the natural logarithm of one plus the yearly review volume from year $t-1$ to year t. Similarly, Δ*Mean rating*$_{i,t}$ and Δ*%Bug*$_{i,t}$ are the yearly differences in *Mean rating* and *%Bug*. We measure firms' financing outcomes in two ways: $I_{Funding}$, a binary indicator that equals one if the startup obtained any equity financing in the subsequent year; and *Log(Total funding)*, the natural logarithm of one plus the total financing amount raised by the startup in the subsequent year. We utilize these two measures to assess the correlation between changes in app performance metrics and the likelihood and the amount of financing.

Table A.14 indicates a significant correlation between app performance metrics and future financing success. Specifically, a 10% increase in review volume is associated with a 0.17 percentage-point increase (i.e., $ln(1+10\%) \times 0.018$, or 0.7% of the sample mean, 23.65 percentage points) in the probability of obtaining financing in the subsequent year and a 0.9% (i.e., $e^{(ln(1+10\%)\times 0.09)} - 1$) increase in the total financing amount secured in that period. Translating our estimated GV treatment effects into these terms, the 115% increase in review volume for GV-backed Android apps reported in column (3) of Table 5 would imply a 1.38 percentage-point increase (i.e., $\ln(1+115\%) \times 0.018$, or 5.8% of the sample mean, 23.65 percentage points) in their likelihood of securing financing in the next year, along with a 7.1% increase (i.e., $e^{(\ln(1+115\%)\times 0.09)} - 1$) in total financing raised.

Similarly, a one-unit increase in the mean rating correlates with a 1.3 percentage-point increase (or 5.5% of the sample mean) in the probability of securing financing in the following year and a 4.3% increase in the total amount of financing. Thus, the 0.45 increase in mean rating shown in column (3) of Table 6 would translate into a 0.6 percentage-point increase (or 2.5% of the sample mean) in the probability of securing financing in the following year, and a 1.9% increase in the total financing amount. In contrast, the number of bug-related reviews does not exhibit a significant direct correlation with financing outcomes.

While the perceived technical quality of an app may not directly translate into better financing outcomes, it might indirectly contribute to firms' financing by boosting future user downloads and usage. This evidence is also consistent with our earlier analysis, which suggests that GV might have taken deliberate actions to delete the negative and bug-related ratings for the apps developed by their investees and that these apps are associated with higher review volume after GV's investments. To examine this possibility, in column 3 of Table A.14, we examine the correlation between changes in the app's rating and bug reports and future app review volume while controlling for current changes in review volume. The estimates suggest that improvements in current rating and bug reports are significantly predictive of future review volume, suggesting higher future user downloads and usage. These correlations are consistent with our conjecture that platform curation of the review data might have contributed to favorable financial outcomes either directly or indirectly by boosting future user demand.

A.6 Startup Exit Outcomes

Table A.15 reports the exit outcomes for the startups in the sample. By the end of 2023, 26.19% of GV-invested startups have achieved a successful exit and 11.90% have successfully completed an initial public offering (IPO).[1] Both of these numbers exceed those for startups associated with the other VCs, though the differences are not statistically significant.

The results in Tables A.13 to A.15 all point to a significant association between GV ownership and the long-term success of app-developing startups. While these results do not account for potential endogeneity, such as GV's superior startup selection skills, they are consistent with our evidence that GV plays a proactive role in boosting its investees' app performance on its own platform and suggest the tangible benefits of their intervention. These tangible benefits further justify GV's motivation to promote its invested apps using its market power in the mobile app market.

A.7 Extending Preferential Treatment to Investment Partners

Evidence in Section 3 indicates that GV-backed apps receive preferential support from the Google Play platform, boosting their visibility and perceived quality. Along these lines, it is plausible that VCs with established connections to GV also derive benefits from the platform, given the existing evidence of the value of VC networks (Hochberg et al., 2007; Sorenson and Stuart, 2001). To examine this possibility, we extend the analysis in Table 2 and Table 4 to apps associated with VCs that are connected with GV through previous syndication (referred to as connected-VC hereafter).

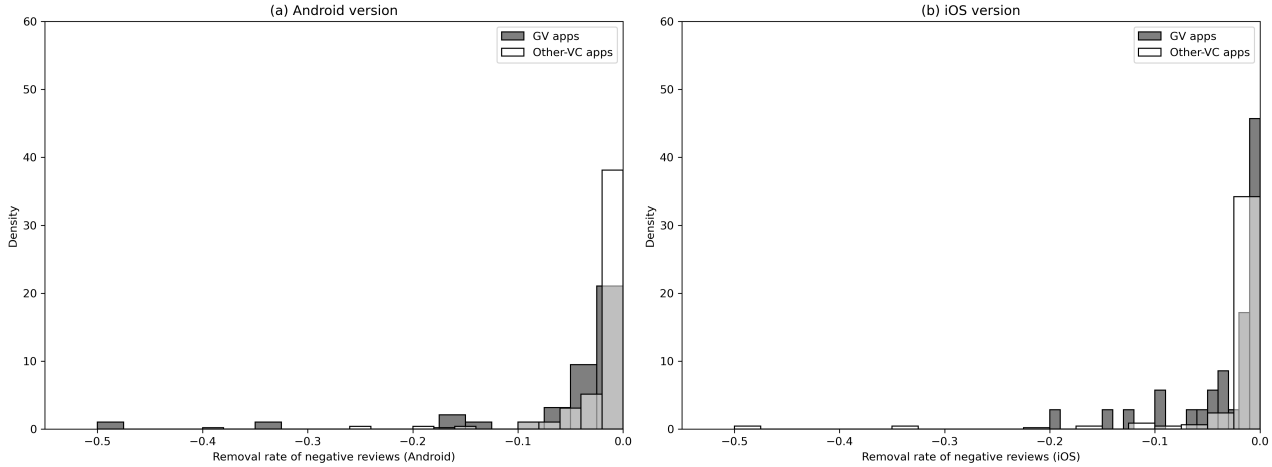
Panel A of Table A.16 presents the difference in app visibility between connected-VC apps and other-VC apps on Google Play, following the methodology in Table 2. Connected-VC apps are recommended as similar apps an average of 15.08 times. Of these recommendations, 7.50 times are from apps within the same genre, while 7.58 times come from apps across different genres. Moreover, these apps are recommended by apps from 2.20 different genres on average. Additionally, connected-VC apps display significantly higher visibility across all these dimensions compared to other-VC apps. However, their advantage in visibility over other-VC apps is smaller in magnitude compared to GV apps, as shown in Table 2.

Panel B of Table A.16 presents the difference in the review removal rate between connected-VC apps and other-VC apps from Google Play, following the methodology outlined in Panel A of Table 4. The results show that 2.69% of negative reviews for connected-VC apps were removed in the second snapshot, compared to 1.75% for other-VC apps. The difference of -0.95% (T = -2.78) is statistically significant at the 1% level, suggesting potential review curation for apps backed by connected VCs. Similarly, connected-VC apps exhibit a 2.06% reduction in the volume of bug reports, compared to a 1.10% reduction for other-VC apps, with the difference statistically significant at the 1% level. However, the removal rates of negative and bug-related reviews for connected-VC apps are less pronounced than those observed for GV apps, as reported in Panel A of Table 4.

These results are collectively consistent with the notion that preferential support from the platform might extend to VCs with established connections to the platform owner, albeit to a lesser degree than that enjoyed by GV apps. This evidence lends additional credence to the benefits of syndication networks and highlights the potential advantages of having strong ties to the platform owner.

[1] Successful exit events include mergers, IPOs, secondary sales, direct public offerings, write-offs, buybacks, and reverse takeovers.

Figure A.1: The distribution of the removal rate of negative reviews



This figure presents the density histogram of the removal rate of negative reviews for each platform. Negative reviews refer to reviews with 1 and 2-star ratings. The removal rate for each app is calculated as the relative change in the total number of negative reviews between the two review data snapshots for the same period from July 2022 to June 2023. Subfigure (a) shows the distribution for 38 Android-version GV apps (black bars) and 244 Android-version apps backed by other VCs (white bars). Subfigure (b) shows the distribution for 35 iOS-version GV apps (black bars) and 186 iOS-version apps backed by other VCs (white bars). *GV app* refers to apps developed by startups backed by GV and without an exit status as of July 2023, ensuring GV's ownership during the data collection period. *Other-VC app* refers to apps developed by startups backed by VCs that have never co-invested with GV.

Table A.1: Percentage of apps with at least 10% removal of negative reviews

	Android apps	iOS apps
GV app	13.16%	8.57%
Other-VC app	3.69%	5.91%
GV app − Other-VC app	9.47%**	2.66%
	(Z = 2.50)	(Z = 0.59)

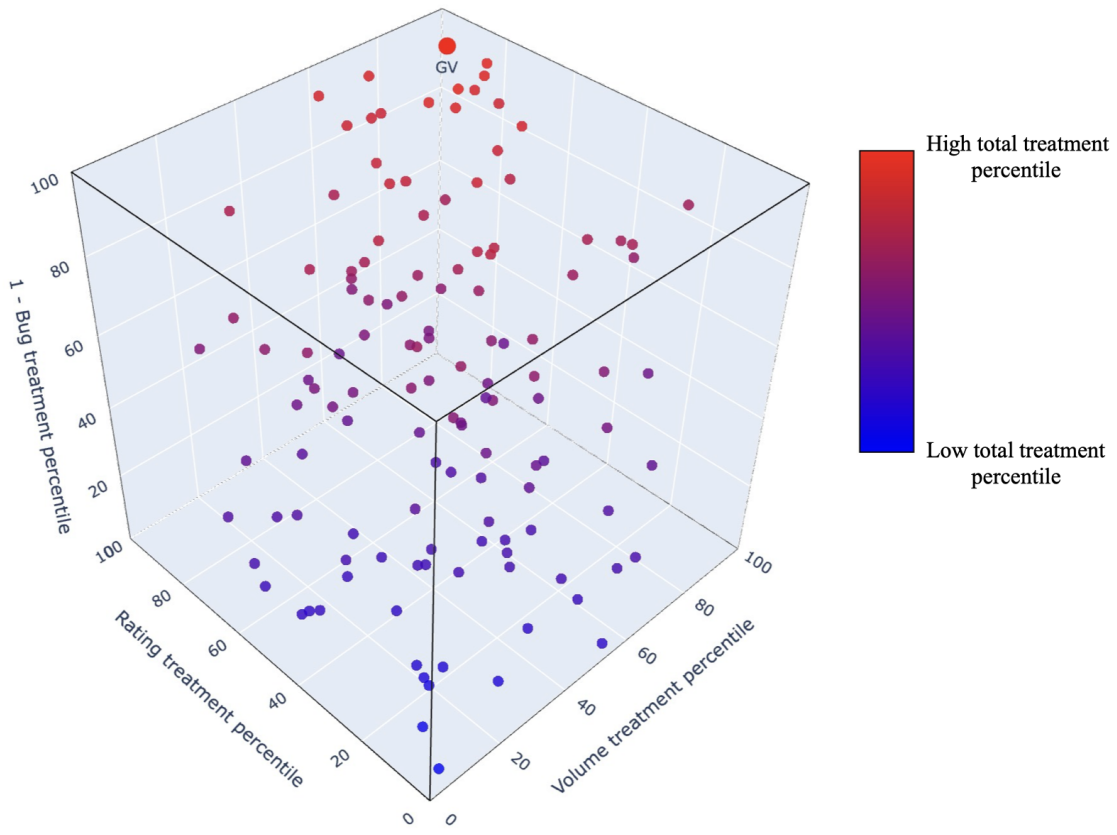
This table shows the percentage of apps with at least 10% of their negative reviews removed between two data snapshots for GV-backed and other-VC-backed apps, across Android and iOS platforms. The data covers reviews posted from July 2022 to June 2023. For Android apps, the snapshots were collected in July 2023 and January 2024; for iOS apps, the snapshots were collected in September 2023 and September 2024. Negative reviews are defined as those with 1- or 2-star ratings. *GV app* refers to apps developed by startups backed by GV and without an exit status as of July 2023, ensuring GV's ownership during the data collection period. *Other-VC app* refers to apps developed by startups backed by VCs that have never co-invested with GV. We report z-statistics for the differences in proportions between the two groups in parentheses. *, ** and *** indicate significance better than 10%, 5%, and 1% respectively.

Table A.2: Content analysis of removed reviews

	Accuracy	Precision	Recall
Training set (other-VC apps)	0.61	0.62	0.57
Validation set (other-VC apps)	0.61	0.62	0.57
Testing set (GV apps)	0.53	0.55	0.36

This table presents the performance of the trained neural network model across different datasets. We construct a balanced dataset of 4,160 removed negative reviews and 4,160 randomly chosen retained negative reviews from the Android reviews of other-VC apps, covering the same period (July 2022 to June 2023). We use 80% of this dataset as a training set to train the model and the remaining 20% as a validation set for out-of-sample testing. During the training stage, we deploy the BERT model to process the review texts from the training set and use the processed outputs as inputs for a 4-layer neural network to test if the model could differentiate between removed and retained reviews based on review content. Then, we conduct separate out-of-sample tests on the validation set from other-VC apps and the testing set from GV apps. The testing set includes 678 negative reviews from GV apps, among which 339 are removed negative reviews and 339 are randomly chosen from the retained negative reviews.

Figure A.2: Comparative ranking of GV treatment effects



This figure plots the percentile ranking of VC treatment effects for all the 128 VCs (including GV) that have invested in a minimum of five startups included in the sample. We estimate the same DiD regressions for each of the 128 VCs, following the specification of column 3 of Tables 5, 6, and 7. The volume treatment effect percentile is based on the ranking of t-statistics for the interaction term *Post* × *Android app*, with *Log(review volume)* as the dependent variable. The rating treatment effect percentile is based on the ranking of t-statistics for the interaction term *Post* × *Android app*, with *Mean rating* as the dependent variable. The bug treatment effect percentile is based on the reverse ranking of t-statistics for the interaction term *Post* × *Android app*, with *%Bug* as the dependent variable. We then calculate the total treatment rank for each VC based on the sum of the three treatment percentiles. Each dot in the graph represents one VC, with the dot at the top corresponding to GV, which ranks the highest overall in terms of treatment effect among all 128 VCs.

Table A.3: Changes in the update frequency after GV investment: Android versus iOS

	Update frequency		
	(1)	(2)	(3)
Android app	-0.005	-0.005	-0.005
	(-0.01)	(-0.01)	(-0.01)
Post	1.802**	0.751	
	(2.25)	(0.77)	
Post × Android app	0.230	0.230	0.230
	(0.37)	(0.36)	(0.37)
App pair FE	N	Y	N
Quarter FE	N	Y	N
App pair × Quarter FE	N	N	Y
Observations	1494	1494	1494
Adjusted R^2	0.025	0.534	0.733

This table presents the estimates of Model (3) using app update frequency as the dependent variable. The Android and iOS versions of the same app constitute an app pair. The sample consists of 1,494 app-quarter observations from January 2010 to June 2023 for 22 app pairs developed by 17 startups that were backed by GV during that period. To ensure a sufficient pre-investment period, an app pair is included in the sample only if both its Android and iOS versions were launched at least two quarters before GV's investment. *Android app* is a binary variable that equals one for the Android version of an app pair, and zero for the iOS version. *Post* is a binary variable that equals one for observations after GV's investment. We include alternative combinations of app pair fixed effects, quarter fixed effects, and app pair × quarter fixed effects in the regressions. We report t-statistics based on standard errors clustered at the app pair level in parentheses. *, ** and *** indicate significance better than 10%, 5%, and 1% respectively.

Table A.4: GV treatment effects and major app updates

	Log(review volume) (1)	Mean rating (2)	%Bug (3)
Android app	0.003	-0.378**	0.034***
	(0.01)	(-2.24)	(2.97)
Lag major update	-0.195	0.115	0.048*
	(-0.52)	(0.51)	(1.85)
Post × Android app	0.826***	0.489**	-0.040***
	(3.17)	(2.58)	(-3.20)
Android app × Lag major update	1.110**	0.276	-0.058
	(2.59)	(0.92)	(-1.16)
Post × Lag major update	0.217	-0.089	-0.035
	(0.43)	(-0.38)	(-1.28)
Post × Android app × Lag major update	-0.726*	-0.260	0.043
	(-1.76)	(-0.85)	(0.86)
App pair × Quarter FE	Y	Y	Y
Observations	1494	1328	1328
Adjusted R^2	0.921	0.440	0.186

This table presents the estimates of Model (3) augmented with a triple interaction between *Post*, *Android app*, and *Lag major update*, using three different measures of app performance as the dependent variable: the natural logarithm of one plus the number of reviews (*Log(review volume)*), the average app rating (*Mean rating*), and the proportion of reviews that contain keywords related to app bugs (*%Bug*). The Android and iOS versions of the same app constitute an app pair. The sample consists of 1,494 app-quarter observations, including 1,328 observations with non-zero reviews, from January 2010 to June 2023 for 22 app pairs developed by 17 startups that were backed by GV during that period. To ensure a sufficient pre-investment period, an app pair is included in the sample only if both its Android and iOS versions were launched at least two quarters before GV's investment. *Android app* is a binary variable that equals one for the Android version of an app, and zero for the iOS version. *Post* is a binary variable that equals one for observations after GV's investment. *Lag major update* is a binary variable equal to one if the app underwent a major version change in the previous quarter, and zero otherwise. We include alternative combinations of app pair fixed effects, quarter fixed effects, and app pair × quarter fixed effects in the regressions. We report t-statistics based on standard errors clustered at the app pair level in parentheses. *, ** and *** indicate significance better than 10%, 5%, and 1% respectively.

Table A.5: Co-investment patterns among GV-backed startups

	# Rounds	Average # co-investors
Startup without apps	692	5.36
Startup with apps	146	4.81
Startup without apps − Startup with apps		0.55**
		(T = 2.29)

This table compares co-investment patterns across GV-backed startups with and without mobile apps matched in our sample. The sample includes all GV-invested financing rounds from 2008 to 2022. *Startup with apps* refers to GV-backed startups whose apps can be matched in our app-level sample, while *Startup without apps* refers to GV-backed startups without matched apps. The table reports the total number of financing rounds and the average number of co-investors per round for each group. We conduct t-tests on the differences in the average number of co-investors and report the t-statistics in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A.6: GV treatment effects and lead-investor status

	Log(review volume) (1)	Mean rating (2)	%Bug (3)
Android app	0.314	-0.414*	0.015
	(0.94)	(-1.78)	(0.91)
Post × Android app	0.500*	0.497**	-0.031
	(1.79)	(2.17)	(-1.64)
Android app × Lead VC	-0.421	0.184	0.022
	(-0.89)	(0.74)	(0.95)
Post × Android app × Lead VC	0.677	-0.107	-0.005
	(1.33)	(-0.38)	(-0.21)
App pair × Quarter FE	Y	Y	Y
Observations	1494	1328	1328
Adjusted R^2	0.919	0.441	0.185

This table presents estimates of Model (3) augmented with a triple interaction that captures GV's lead-VC status. We test on three different measures of app performance as the dependent variables: the natural logarithm of one plus the number of reviews (*Log(review volume)*), the average app rating (*Mean rating*), and the proportion of reviews that contain keywords related to app bugs (*%Bug*). The Android and iOS versions of the same app constitute an app pair. The sample consists of 1,494 app-quarter observations, including 1,328 observations with non-zero reviews, from January 2010 to June 2023 for 22 app pairs developed by 17 startups that were backed by GV during that period. To ensure a sufficient pre-investment period, an app pair is included in the sample only if both its Android and iOS versions were launched at least two quarters before GV's investment. *Android app* is a binary variable that equals one for the Android version of an app, and zero for the iOS version. *Post* is a binary variable that equals one for observations after GV's investment. *Lead VC* is a binary variable equal to one if GV served as the lead investor in the startup's first financing round in which GV participated. Among the 17 sample startups in the DiD sample, GV acted as the lead VC for 9 of them. We include app pair × quarter fixed effects in the regressions. We report t-statistics based on standard errors clustered at the app pair level in parentheses. *, ** and *** indicate significance better than 10%, 5%, and 1% respectively.

Table A.7: Review removals by startup financing stage

	# Obs	% Change in negative review vol	% Change in neutral review vol	% Change in positive review vol	% Change in bug-related review vol	Change in mean rating
Early stage	7	-7.42%	-2.61%	-1.35%	-12.43%	0.041
Expansion stage	14	-4.46%	-1.98%	-1.50%	-3.06%	0.028
Later stage	17	-4.92%	-1.57%	-1.06%	-2.88%	0.007
Early stage − Later stage		-2.51% (T = -0.61)	-1.05% (T = -0.59)	-0.29% (T = -0.27)	-9.55% (T = -1.49)	0.034 (T = 1.87)

This table compares review removal patterns across startups at different financing stages among 38 GV apps. The data covers reviews posted between July 2022 and June 2023 and follows the snapshot comparison methodology in Table 4. For Android apps, the two snapshots were collected in July 2023 and January 2024. Review types are categorized into negative (1–2 stars), neutral (3–4 stars), positive (5 stars), and bug-related reviews. Financing stage (early stage, expansion stage and later stage) is defined based on the startup's most recent financing round before the end of 2022, as classified in the LSEG Workspace database. We report t-statistics for the differences in the review removal rate and mean rating in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A.8: App visibility and early-stage status

	Overall visibility	Within-category visibility	Cross-category visibility	# Visible categories
	(1)	(2)	(3)	(4)
GV	0.890**	0.305	1.582***	1.129***
	(2.29)	(0.65)	(3.21)	(2.76)
Early stage	-0.814*	-0.792	-0.679	-0.569
	(-1.83)	(-1.46)	(-1.49)	(-1.33)
GV × Early stage	1.461**	0.923	1.196*	0.861
	(2.25)	(1.19)	(1.73)	(1.30)
App age	-0.029	0.084	-0.121**	-0.064
	(-0.59)	(1.36)	(-2.51)	(-1.40)
Log(# downloads)	0.301***	0.270***	0.339***	0.294***
	(6.71)	(4.68)	(5.81)	(5.84)
Mean rating	0.579**	0.697**	0.548**	0.468**
	(2.51)	(2.10)	(2.27)	(2.32)
Log(Total funding)	-0.018	0.034	-0.066*	-0.029
	(-0.62)	(0.85)	(-1.92)	(-1.03)
App category FE	Y	Y	Y	Y
Firm industry FE	Y	Y	Y	Y
Observations	287	287	276	287
Pseudo R^2	0.612	0.551	0.678	0.537

This table presents the estimates of cross-sectional Poisson regressions of app visibility augmented with financing-stage heterogeneity. The dependent variable is the visibility of an Android app measured in four different ways: total number of recommendations as "similar app" (*Overall visibility*), number of recommendations as "similar app" from apps of the same genre (*Within-category visibility*), number of recommendations as "similar app" from apps of different genres (*Cross-category visibility*), and the number of different genres where the app shows up as a "similar app" (*# Visible categories*). *GV* is a binary variable that equals one for apps developed by startups invested by GV. *Early stage* is a binary variable that equals one if the startup's most recent financing round before the end of 2022 is classified as early-stage financing in the LSEG Workspace database. We include app category (genre) and firm industry fixed effects in the regressions. We report t-statistics based on robust standard errors in parentheses. *, ** and *** indicate significance better than 10%, 5%, and 1% respectively.

Table A.9: GV treatment effects and early-stage status

	Log(review volume) (1)	Mean rating (2)	%Bug (3)
Android app	0.033	-0.412*	0.033*
	(0.11)	(-2.04)	(1.89)
Post × Android app	0.890**	0.525**	-0.043**
	(2.69)	(2.58)	(-2.40)
Android app × Early stage	0.377	0.219	-0.026
	(0.78)	(0.90)	(-1.31)
Post × Android app × Early stage	-0.513	-0.107	0.032
	(-0.84)	(-0.46)	(1.60)
App pair × Quarter FE	Y	Y	Y
Observations	1494	1328	1328
Adjusted R^2	0.918	0.441	0.182

This table presents the estimates of Model (3) augmented with financing-stage heterogeneity. The Android and iOS versions of the same app constitute an app pair. The sample consists of 1,494 app-quarter observations, including 1,328 observations with non-zero reviews, from January 2010 to June 2023 for 22 app pairs developed by 17 startups that were backed by GV during that period. To ensure a sufficient pre-investment period, an app pair is included in the sample only if both its Android and iOS versions were launched at least two quarters before GV's investment. *Android app* is a binary variable that equals one for the Android version of an app, and zero for the iOS version. *Post* is a binary variable that equals one for observations after GV's investment. *Early stage* is a binary variable that equals one if the current period occurs before the startup's most recent early-stage financing round, as classified in the LSEG Workspace database. Out of the 17 sample startups, 8 were at an early stage at some point, meaning GV invested in them while they were still in the early-stage phase. We include app pair × quarter fixed effects in the regressions. We report t-statistics based on standard errors clustered at the app pair level in parentheses. *, ** and *** indicate significance better than 10%, 5%, and 1% respectively.

Table A.10: Review removals by startup exit status

	# Obs	% Change in negative review vol	% Change in neutral review vol	% Change in positive review vol	% Change in bug-related review vol	Change in mean rating
GV app	38	-5.21%	-1.91%	-1.28%	-4.71%	0.021
Exited GV app	15	-3.68%	-2.22%	-1.45%	-3.12%	0.017
GV app − Exited GV app		-1.53% (T = -0.80)	0.31% (T = 0.42)	0.18% (T = 0.41)	-1.59% (T = -0.87)	0.004 (T = 0.31)

This table compares review removal patterns between apps developed by startups currently backed by GV and those previously backed by GV that had achieved exit status by the end of 2022. The data covers reviews posted between July 2022 and June 2023 and follows the snapshot comparison methodology in Table 4. For Android apps, the two snapshots were collected in July 2023 and January 2024. Review types are categorized into negative (1–2 stars), neutral (3–4 stars), positive (5 stars), and bug-related reviews. *GV app* refers to apps developed by startups invested by GV that have not achieved exit status by the end of 2022, ensuring GV's ownership during the data collection period. *Exited GV app* refers to apps developed by startups previously backed by GV but that have achieved exit status by the end of 2022. We report t-statistics for the differences in the review removal rate and mean rating in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A.11: Robustness: Event window restricted to 12 quarters around GV investment

	Log(review volume) (1)	Mean rating (2)	%Bug (3)
Android app	0.202	-0.324**	0.024*
	(0.82)	(-2.36)	(1.74)
Post × Android app	0.325*	0.284**	-0.031*
	(1.76)	(2.42)	(-2.06)
App pair × Quarter FE	Y	Y	Y
Observations	910	768	768
Adjusted R^2	0.924	0.378	0.126

This table presents the estimates of Model (3) using three different measures of app performance as the dependent variables: the natural logarithm of one plus the number of reviews (*Log(review volume)*), the average app rating (*Mean rating*), and the proportion of reviews that contain keywords related to app bugs (*%Bug*). The Android and iOS versions of the same app constitute an app pair. The sample is restricted to observations within 12 quarters before and 12 quarters after GV's investment for each app pair. *Android app* is a binary variable that equals one for the Android version of an app, and zero for the iOS version. *Post* is a binary variable that equals one for observations after GV's investment. We include app pair × quarter fixed effects in the regressions. We report t-statistics based on standard errors clustered at the app pair level in parentheses. *, ** and *** indicate significance better than 10%, 5%, and 1% respectively.

Table A.12: Robustness: Two-stage difference-in-differences estimation

	Log(review volume) (1)	Mean rating (2)	%Bug (3)
Post × Android app	0.920***	0.077	-0.008
	(Z = 5.22)	(Z = 0.93)	(Z = -1.55)
App pair × Quarter FE	Y	Y	Y
Observations	1494	1392	1392

This table presents robustness estimates using the Two-stage difference-in-differences estimator proposed by Gardner (2022). The dependent variables are the natural logarithm of one plus the number of reviews (*Log(review volume)*), the average app rating (*Mean rating*), and the proportion of reviews that contain keywords related to app bugs (*%Bug*). The Android and iOS versions of the same app constitute an app pair. *Android app* is a binary variable that equals one for the Android version of an app, and zero for the iOS version. *Post* is a binary variable that equals one for observations after GV's investment. We include app pair × quarter fixed effects in the estimation. We report z-statistics based on standard errors clustered at the app pair level in parentheses. *, ** and *** indicate significance better than 10%, 5%, and 1% respectively.

Table A.13: Financing outcomes across VC types

	# startups	Average funding rounds	Average total funding	Average funding per round
GV	42	5.95	639.46	72.86
Other VC	197	2.68	30.29	11.16
GV − Other VC		3.28*** (T = 5.78)	609.16** (T = 2.58)	61.70*** (T = 3.54)

This table presents the financing outcomes for startups affiliated with GV and other VCs that have never engaged in co-investments with GV during the sample period. *Funding rounds* is the cumulative number of funding rounds secured by the startups during the sample period. *Total funding* (in millions) is the cumulative amount of financing raised by the startups throughout the sample period. *Funding per round* (in millions) is the average amount of financing the startups raise per funding round. We report t-statistics for the differences in the financing outcomes between the two groups in parentheses. *, ** and *** indicate significance better than 10%, 5%, and 1% respectively.

Table A.14: App performance and financing outcomes

	One-year-ahead $I_{Funding}$	One-year-ahead Log(Total funding)	One-year-ahead Log(review volume)
	(1)	(2)	(3)
ΔLog(review volume)	0.018**	0.090***	0.170***
	(2.39)	(3.43)	(8.35)
ΔMean rating	0.013*	0.042*	0.044***
	(1.65)	(1.68)	(2.77)
Δ%Bug	0.045	0.162	-0.233**
	(0.81)	(0.91)	(-1.97)
Firm FE	Y	Y	Y
Year FE	Y	Y	Y
Observations	4947	4947	4032
Adjusted R^2	0.174	0.214	0.910

This table reports the estimates of startup firm-year panel regressions. The sample includes 1,024 startups with VC financing and at least one Android app. In columns 1 and 2, the dependent variables are measures of financing outcomes, including a binary variable that equals one if the sample startup obtained financing in the subsequent year (*One-year-ahead* $I_{Funding}$) and the natural logarithm of one plus the total amount of financing raised by the startup in the subsequent year (*One-year-ahead Log(Total funding)*). In column 3, the dependent variable is the natural logarithm of one plus the total review volume in the subsequent year (*One-year-ahead Log(review volume)*). Δ*Log(review volume)* is the yearly difference in the natural logarithm of one plus the yearly review volume. Similarly, Δ*Mean rating* is the yearly difference in *Mean rating* calculated from the review ratings. Δ*%Bug* is the yearly difference in *%Bug*. For each startup, we measure the app-level metrics in the regressions based on its most downloaded app. We include firm and year fixed effects in the regressions. We report t-statistics based on standard errors with startup clustering in parentheses. *, ** and *** indicate significance better than 10%, 5%, and 1% respectively.

Table A.15: Startup exit status across VC types

	# startups	# exits	# IPO	% IPO	% Exit
GV	42	11	5	11.90%	26.19%
Other VC	197	42	11	5.58%	21.32%
GV − Other VC				6.32% (Z = 1.49)	4.87% (Z = 0.69)

This table presents the rate of successful exit for startups backed by GV and other VCs that have no syndication connection with GV. *% Exit* indicates the percentage of startups that have achieved a successful exit by the end of the sample period. Successful exit events include mergers, IPOs, secondary sales, direct public offerings, write-offs, buybacks, and reverse takeovers. *% IPO* is the percentage of startups that have successfully completed an initial public offering by the end of the sample period. We report z-statistics for the differences in proportions between the two groups in parentheses. *, ** and *** indicate significance better than 10%, 5%, and 1% respectively.

Table A.16: App visibility and the review removal rate for Android apps backed by connected VCs

Panel A: App visibility					
	# Apps	Overall visibility	Within-category visibility	Cross-category visibility	# Visible categories
Connected-VC app	1092	15.08	7.50	7.58	2.20
Other-VC app	287	6.85	3.00	3.85	1.18
Connected-VC app − Other-VC app		8.23***	4.50***	3.73***	1.02***
		(T = 5.33)	(T = 5.49)	(T = 3.74)	(T = 4.33)

Panel B: Changes in review volume						
	# Apps	% Change in negative review vol	% Change in neutral review vol	% Change in positive review vol	% Change in bug-related review vol	Change in mean rating
Connected-VC app	958	-2.69%	-2.47%	-1.82%	-2.06%	0.000
Other-VC app	244	-1.75%	-1.64%	-0.97%	-1.10%	0.004
Connected-VC − Other-VC app		-0.95%***	-0.83%*	-0.85%***	-0.96%***	-0.004
		(T = -2.78)	(T = -1.93)	(T = -3.11)	(T = -2.85)	(T = -1.22)

This table presents the differences in app visibility and the review removal rate between apps backed by connected VCs and those backed by other VCs. *Connected-VC app* refers to apps developed by startups invested by VCs that are connected with GV through syndication. *Other-VC app* refers to the apps developed by startups invested by VCs that have never co-invested with GV. Panel A of Table A.16 presents the difference in app visibility between connected-VC apps and other-VC apps, following the methodology in Table 2. The visibility of an Android app is measured in four ways: total number of appearances as "similar app" (*Overall visibility*), number of appearances as "similar app" for apps of the same genre (*Within-category visibility*), number of appearances as "similar app" for apps of different genres (*Cross-category visibility*), and the number of different genres where the app shows up as a "similar app" (*# Visible categories*). We conduct t-tests on the differences in visibility between connected-VC apps and other-VC apps and present the t-statistics in parentheses. *, ** and *** indicate significance better than 10%, 5%, and 1% respectively. Panel B of Table A.16 presents the difference in the review removal rate between connected-VC apps and other-VC apps from Google Play, following the methodology in Panel A of Table 4. We separately count four types of reviews: negative reviews (1 and 2-star ratings), neutral reviews (3 and 4-star ratings), positive reviews (5-star ratings), and bug-related reviews (reviews containing keywords related to app bugs). We report t-statistics for the differences in the review removal rate in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.