Spring 2025
Econ 452: Econometrics

The Run Support Riddle: The Influence of Team Run Production on Pitcher Win Totals

Ibrahim Kaba
5/30/25

Introduction

In today's world sports are everywhere. Not only that, sports also play a meaningful role in a country's economy. In fact, they "are well recognized as a significant contributor to the Gross Domestic Product, consumer expenditures and employment opportunities" (Chen, Gunter, Zhang, 429) It's very hard to picture life without them. There is one sport in particular that deserves considerable attention. The only sport on the planet where you can fail seventy percent of the time and still be considered great and earn very handsome wages. The best of the best baseball players compete in the North American league called Major League Baseball (MLB). To say the least there was a period in history when baseball was judged as an American pastime and treated more like a religion (Chen, Gunter, Zhang, 430). Times have changed and so has the MLB. One of the many ways the league has evolved is through its utilization of statistics and

data analytics to enable improved " decision-making in player acquisitions", analysis in-game data , real-time decisions, evaluation of player performance and much more (Fiasconaro). The game's reliance on metrics is undeniable. The metric that the study is concerned with is the pitcher's win-loss record. A pitcher's win-loss record represents the foundation of individual success in a very precious position . However, the measure is often disputed by millions around the world for its failure to capture an accurate representation of a pitcher's performance. For example, a pitcher may have an outstanding day by limiting runs, forcing the opposition into weak contact, and succeeding in high pressure situations only to suffer a loss if their team's offense struggles. The opposite is also possible as well. A mediocre pitching performance could also result in a win by brilliant offensive effort. More often than not both types of situations occur. A critical question arises:

To what extent does Run Support per Game Started (RS/GS) influence a pitcher's win total, after controlling for pitching durability and effectiveness as measured by IP/GS and QS%?

This study aims to evaluate the causal impact of offensive support on pitcher win totals while accounting those core indicators of pitching performance publicly available . The importance of the research question cannot be stressed enough because at the end of the day the MLB is a business. A business needs to make money to survive and if your team is not producing results then you will not last very long in the league. And so with that being said, win totals have the capacity to influence contract negotiations, scouting reports, professional bettors, and much more. In a nutshell the insight derived from the research contributes to a more accurate and

economically rational assessment of a pitcher's performance in professional baseball which is extremely valuable given the context.

Literature Review

"Win Shares," brought to you by Bill James and Jim Henzler, offers a multifaceted approach to evaluating baseball player contributions that's never been seen before. Rather than controlling for important statistics like ERA (Earned Run Average) or RBIs (Runs Batted In), the new system combines offensive, defensive, and pitching performances into a more complete framework of player value (James & Henzler, 16-100). The author's main discussion centers around the question : "How can we fairly attribute wins to individual players in a team-oriented sport?" (James & Henzler, 3) Most importantly their results were quite remarkable. In particular they found that the impact of external factors like run support on win totals was significant. For example, they report that "for every additional run of support per game, a pitcher's win total could increase by approximately 0.4 wins over a 30-start season—holding other variables constant"(James & Henzler, 123). What exactly does this mean? In simpler terms their findings point to the fact that win totals are not a reflection of great pitching but more intertwined with overall team strength and balance. The reliability of the study is demonstrated through its meticulous data collection and proper model calibration as a result the study has strong internal validity. With regard to external validity the model might not work as well when applied to different time periods or leagues.

Next the paper "Pay and Performance in Major League Baseball," presents another vital contribution by investigating how player salaries correlate with traditional and advanced

performance metrics. The main hypothesis involves the relationship market compensation and individual contributions, for pitchers and hitters. The author, Gerald W. Scully, cunningly, decided to make use of multiple multivariate regression models and interestingly enough he notes “in this study, the lifetime strikeout-to-walk ratio SW is employed as the single best measure of pitching performance”(Scully, 925). After Scully runs the log linear regressions on player salaries, he uncovers the most deterministic performance metrics for pitchers are strikeout-to-walk ratio, innings pitched, and career length (Scully, 924-927). To say the bare minimum, the results are jaw dropping as Scully instigate a shift in how pitcher effectiveness is seen. The regression results are very impressive with “78 to 81 percent of the variation in ballplayers accounted for” (Scully, 927).Win totals may be irrelevant according to Scully in this context , his research still acknowledges their historical value. The internal validity in the study is rather robust due to its empirical excellence and great data. However and similar to the last study, external validity should not be overlooked in regard to specific factors such as team budget constraints and varying front office strategies.

It’s important to recognize the excellence of both studies. They both strategically assess the connection between individual performance and contextual variables carefully. Evidently the following research attempts to offer a more precise estimation of the determinants in the evaluation of a pitcher’s success compared to a team’s offensive strength.

Data Description

The data source used is the 2024 Major League Baseball starting pitching statistics, published by Baseball-Reference.com. This dataset provides extensive statistics for all starting pitchers across MLB teams for the 2024 season. Since each observation corresponds to a unique pitcher's individual numbers for the 2024 season, the dataset qualifies as a cross-sectional dataset rather than a panel dataset. Moreover the following variables will be used in the study:

- **Wins in games started (Wgs):** The number of games credited as a win to the starting pitcher.
- **Innings Pitched per Game Started (IP/GS):** Reflects pitcher durability and the ability to work deep into games.
- **Quality Start Percentage (QS%):** Percentage of games where the pitcher went at least six innings and allowed three or fewer earned runs.
- **Run Support per Game Started (RS/GS):** The average number of runs scored by the pitcher's team during their starts, representing offensive support.

Table 1. Descriptive Statistics

	Mean	Standard Deviation	Minimum value	Maximum value	Observations
Wgs	3.85	4.55	0.00	18.00	370
RS/GS	4.41	1.64	0.00	12.00	370
QS%	0.23	0.22	0.00	1.00	370
IP/GS	4.43	1.45	0.30	6.70	370

In table 1 , the notable numbers that jump out are the starting pitchers' win average was 3.85 wins in games started, with a standard deviation of 4.55. The average run support per start was 4.41 runs, while pitchers recorded a mean quality start percentage 23 percent . On average, pitchers threw 4.43 innings per game started. The dataset includes 370 observations, capturing considerable variability in pitching performance and offensive support.

Basic Regression (Scatterplots + Simple OLS)

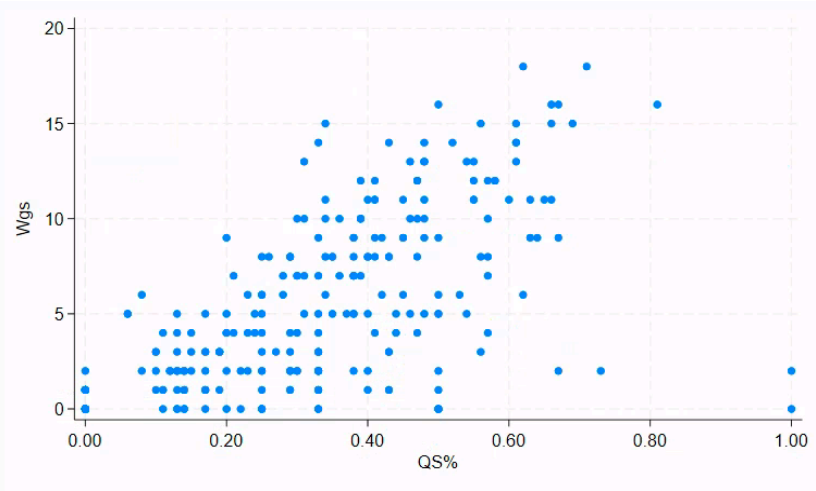


Fig. 1. Wins in Games Started by Quality Start Percentage. *(Stata, 2025).*

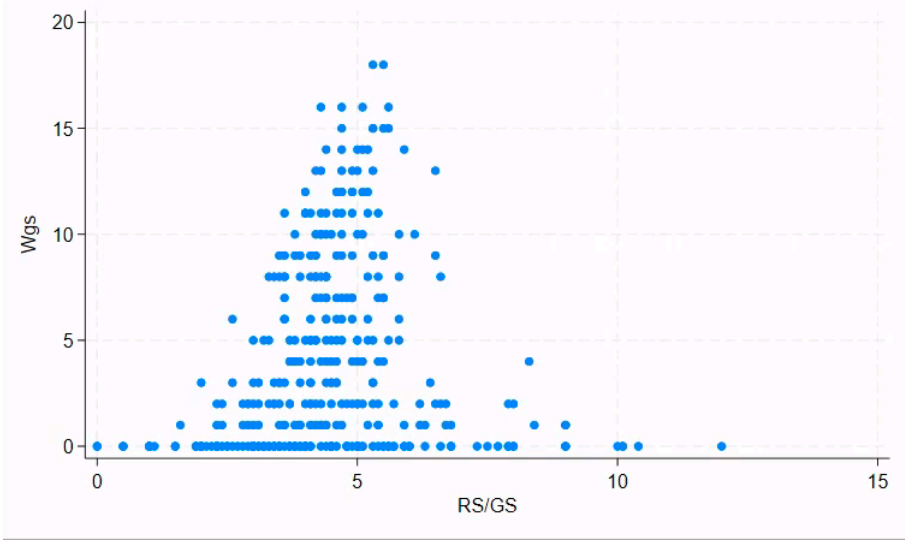


Fig. 2. Wins in Games Started by Run Support per Game Started.. *(Stata, 2025).*

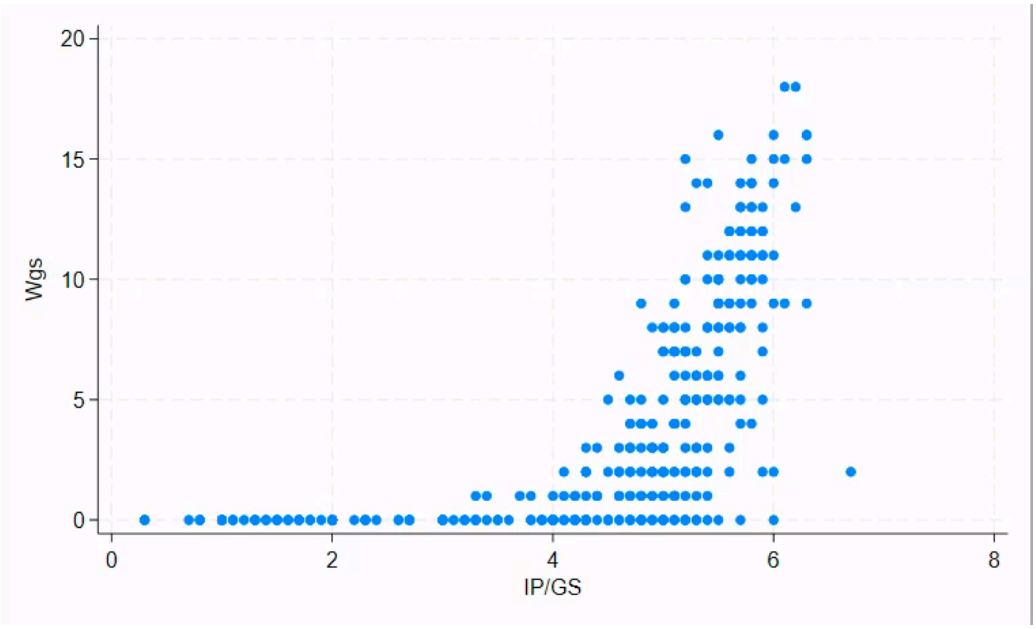


Fig. 3. Wins in Games Started by Innings Pitched per Game Started. *(Stata, 2025).*

Scatterplot Analysis

In *Figure 1* the scatterplot displays the relationship between Wgs and QS%. There is a positive trend, indicating that higher QS% then there will be more wins. Pitchers who deliver more quality starts seem to be more consistent in helping their teams secure the win. Next in *Figure 2* the graph focuses on the relationship between Wgs against RS/GS. There appears to be a positive trend but it is rather weak. Most importantly the relationship flattens at higher levels of RS/GS. This may suggest the concept of diminishing returns in that some pitchers receive strong run support but fail to secure wins, potentially due to poor pitching. Nevertheless while offensive backing matters, it doesn't tell the whole story. Lastly *Figure 3* the scatterplot depicts a strong positive trend between IP/GS and Wgs. In other words, pitchers who go deeper into games tend to acquire more wins. Staying in longer means you are having greater influence over outcomes.

OLS Equation

$$\mathrm{Wgs}_i = \beta_0 + \beta_1 \mathrm{RS/GS}_i + \beta_2 \mathrm{IP/GS}_i + \beta_3 \mathrm{QS\%}_i + \epsilon_i$$

Table 2. Simple OLS Regression Results

	(1) Wgs
RS/GS	0.345***
	(0.102)
QS%	12.94***
	(1.630)
IP/GS	0.536***
	(0.137)
Constant	-3.021***
	(0.649)
Observations	370

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Endogeneity

As mentioned previously, our analysis uses pitcher-level data from the 2024 Major League Baseball season to examine the extent to which Run Support per Game Started (RS/GS) influences a pitcher's win total, while controlling for pitching durability (IP/GS) and effectiveness (QS%). Although the dataset is structured consistently and observations are treated as independent, there are still some issues. In order to validate our OLS estimates potential endogeneity issues must be addressed. One key concern is omitted variable bias. In the study the regression controls for innings pitched per start and quality start percentage, unfortunately there are other unobserved factors that most likely contribute to a pitcher's win total. For example, bullpen quality, bullpen ERA, managerial decisions, defensive support, and opponent strength are not accounted for in our model but could significantly affect win outcomes. The exclusion of these variables could bias our estimates of RS/GS by attributing its effect to unmeasured influences.

A second potential issue is simultaneous causality bias. While our model treats RS/GS as exogenous, it's very much possible that a pitcher performing well (e.g., by pitching deeper into games or limiting runs) may indirectly influence the amount of offensive support they receive. For instance if a pitch is keeping their team in closer games, this could most definitely incite more aggressive offensive strategies. And so, pitching performance could simultaneously affect and be affected by RS/GS, violating the assumption of exogeneity. To mitigate these concerns, we employ the following models:

Reduced Model: While susceptible to omitted variable bias due to excluding IP/GS and QS%, the inflation of RS/GS's coefficient here alerts us to the bias introduced when performance variables are ignored, underlining the necessity of control variables.

Log-Linear Model: The use of this model helps to mitigate heteroskedasticity and normalizes the distribution of win totals, allowing for more stable variance and a more effective inference. It also addresses bias by reducing distortion in estimates that arise from skewed win data, especially when analyzing marginal effects across a wide range.

Interaction Model: The third method confronts unobserved heterogeneity and potential omitted interaction effects by modeling how RS/GS's impact varies with QS%. The result of this method allows for the isolation of the conditional effect of run support.

Results

Table 3.

Comparison of Regression Results

	(1) Wgs	(2) Wgs	(3) ln_Wgs	(4) Wgs
RS/GS	0.345***	0.372***	0.076***	0.073
	(0.102)	(0.140)	(0.023)	(0.102)
QS%	12.945***		2.450***	2.970
	(1.630)		(0.331)	(2.072)
IP/GS	0.536***		0.228***	0.499***
	(0.137)		(0.031)	(0.146)
RSxQS				2.265***
				(0.417)
Constant	-3.021***	2.211***	-0.804***	-1.674**
	(0.649)	(0.579)	(0.145)	(0.710)
Observations	370	370	370	370

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The Reduced model regression of Wgs and RS/GS demonstrates RS/GS has a significant positive effect. However, as mentioned earlier there is a significant risk of omitted variable bias. The coefficient on RS/GS is inflated when IP/GS and QS% are not included, reinforcing the necessity of other performance metrics. On the other hand the Log-Linear model coefficients reflect percentage changes. The RS/GS coefficient remains significant and positive but from *Fig.2* it's evident that the magnitude eventually diminishes. This means that the marginal benefit of run support decreases as total wins rise. Hence the conclusion must be diminishing returns, a key finding. Be that as it may, the insight from the Log-Linear model supports the use of log transformations in baseball economics when dealing with count-based outcomes like win totals. The best was saved for last. In the interaction model regression, there is an examination of whether RS/GS's effect varies with QS%. Results point to moderate interaction, implying that RS/GS is more valuable for pitchers with higher QS%. Nevertheless this finding supports the idea that run support amplifies its effect when a pitcher is already performing well, rather than

compensating for poor performance. The interaction term adds another layer of depth to our understanding of the dynamics between team support and individual skill. The robustness of the model cannot be overlooked. Generally the interaction terms in this context help control for unobservable heterogeneity in pitcher ability, and enhance the explanatory power of our estimates. In summary these models make a valiant attempt to collectively provide a comprehensive view of how RS/GS interacts with core performance metrics. One of the major takeaways is there is a necessity to control for multiple statistics associated with a pitcher's performance to obtain unbiased and meaningful coefficients.

Discussion/Conclusion

Let’s put it all together. The results confirm that both run support and individual performance critically influence win totals. RS/GS is not a one man army in the sense that it cannot explain outcomes independently. Regardless, RS/GS is a statistically robust predictor when positioned with QS% and IP/GS. What about the regression? The results from the regression paint a vivid picture in that even a modest increase in run support substantially improves a pitcher’s chances of recording wins. Next the large magnitude of the QS% coefficient emphasizes the value of consistent, high-quality outings across the entire season. IP/GS further highlights that staying in games longer not only affects direct outcomes but also likely boosts other external factors like team confidence but more analysis will need to be done for certainty. These findings are very discerning but additional analysis could make this paper highly profitable. In the event that you took apart RS/GS by innings could reveal when run support is most effective. In addition the integration bullpen ERA and fielding independent metrics could

provide a more comprehensive picture of win determinants. If additional time and resources were available, we would obtain game-level datasets that capture bullpen ERA, defensive errors, managerial decisions or other variables that reduce endogeneity. Like use of state-of-the-art tracking technology, Statcast would make it possible for next level collection and analysis of a tremendous amount of baseball data (MLB.com). The future is so bright.

References & Appendix

James, Bill, and Jim Henzler. Win Shares. STATS Publishing, 2002

Scully, Gerald W. "Pay and Performance in Major League Baseball." The American Economic Review, vol. 64, no. 6, 1974, pp. 915–30. JSTOR, http://www.jstor.org/stable/1815242. Accessed 28 May 2025.

Baseball Reference. "2024 MLB Starter Pitching Stats." https://www.baseball-reference.com/leagues/majors/2024-starter-pitching.shtml.

OpenAI. ChatGPT. 27 May 2025, https://chat.openai.com/. Accessed 27 May 2025. Used for guidance on statistical analysis and regression modeling in Stata for an econometrics research project.

Chen, Ke, et al. “How Global Is U.S. Major League Baseball? A Historical and Geographic Perspective.” *GeoJournal*, vol. 77, no. 3, 2012, pp. 429–44. *JSTOR*, http://www.jstor.org/stable/23254353. Accessed 30 May 2025.

Fiasconaro, Catherine. "What Do Statistics in Major League Baseball & Production Factory Environments Have in Common?" *LinkedIn*, 14 Oct. 2024, www.linkedin.com/pulse/what-do-statistics-major-league-baseball-production-have-fiasconaro-unhme.

Major League Baseball. “Statcast Glossary.” *MLB.com*, https://www.mlb.com/glossary/statcast. Accessed 31 May 2025.

Log File

---

name: <unnamed>

log: C:\Users\Ibrahim.Kaba001\OneDrive - University of Massachusetts Boston\Econ452\final project\finalpro_KABA_KA

> BA.log

log type: text

opened on: 28 May 2025, 02:16:23

. import excel "Econ 452 final mlb start player pitching 2024.xlsx",firstrow clear

(37 vars, 370 obs)

. ssc install estout, replace

checking estout consistency and verifying not already installed...

all files already exist and are up to date.

. ***provide a summary statistic

.

. describe

Contains data

Observations: 370

Variables: 37

Variable Storage Display Value

name type format label Variable label

Rk int %10.0g Rk

Name str25 %25s Name

Age byte %10.0g Age

Tm str3 %9s Tm

IP double %10.0g IP

G byte %10.0g G

GS byte %10.0g GS

Wgs byte %10.0g Wgs

Lgs byte %10.0g Lgs

ND byte %10.0g ND

Wchp byte %10.0g Wchp

Ltuf	byte	%10.0g	Ltuf
Wtm	byte	%10.0g	Wtm
Ltm	byte	%10.0g	Ltm▼
tmWL	double	%10.0g	tmW-L%
Wlst	byte	%10.0g	Wlst
Lsv	byte	%10.0g	Lsv
CG	byte	%10.0g	CG
SHO	byte	%10.0g	SHO
QS	byte	%10.0g	QS
U	double	%4.2f	QS%
GmScA	double	%10.0g	GmScA
Best	byte	%10.0g	Best
Wrst	byte	%10.0g	Wrst
BQR	byte	%10.0g	BQR
BQS	byte	%10.0g	BQS
sDR	byte	%10.0g	sDR
lDR	byte	%10.0g	lDR
RSGS	double	%10.0g	RS/GS
RSIP	double	%10.0g	RS/IP
IPGS	double	%10.0g	IP/GS
PitGS	byte	%10.0g	Pit/GS
AG	byte	%10.0g	<80
AH	byte	%10.0g	80-99
AI	byte	%10.0g	100-119
AJ	byte	%10.0g	≥120
Max	int	%10.0g	Max

Sorted by:

Note: Dataset has changed since last saved.

.

.

. estpost summarize Wgs RSGS U IPGS

| e(count) e(sum_w) e(mean) e(Var) e(sd) e(min) e(max) e(sum)

---+---

Wgs | 370 370 3.848649 20.74668 4.554852 0 18 1424

RSGS | 370 370 4.406216 2.705571 1.644862 0 12 1630.3

U | 370 370 .2299459 .0486645 .2206003 0 1 85.08

IPGS | 370 370 4.428919 2.089487 1.445506 .3 6.7 1638.7

. esttab using summary_stats.rtf, replace ///

> cells("mean(fmt(%9.2f)) sd(fmt(%9.2f)) min(fmt(%9.2f)) max(fmt(%9.2f)) count(fmt(%9.0g))") ///

> collabels("Mean" "Standard Deviation" "Minimum value" "Maximum value" "Observations") ///

> label noobs nonumber nomtitle

(output written to summary_stats.rtf)

.

. **create a scatterplot of your x and y variables

. scatter Wgs U

. scatter Wgs RSGS

. scatter Wgs IPGS

.

.

.

. ***run a simple ols of your y on x (control variables will be ipgs & qs%)

. regress Wgs RSGS U IPGS, robust

Linear regression Number of obs = 370

F(3, 366) = 174.44

Prob > F = 0.0000

R-squared = 0.5940

Root MSE = 2.914

| Robust

Wgs | Coefficient std. err. t P>|t| [95% conf. interval]

RSGS | .3447796 .1017468 3.39 0.001 .1446978 .5448613

U | 12.94481 1.629888 7.94 0.000 9.739688 16.14993

IPGS | .5360403 .1374277 3.90 0.000 .2657933 .8062873

_cons | -3.02121 .6486063 -4.66 0.000 -4.296673 -1.745747

. esttab, se label star(* 0.10 ** 0.05 *** 0.01) ///

```
> title("Figure 1. Impact of Run Support on Wins") ///
> mtitle("Wins in Games Started") ///
> alignment(D) ///
> nogap
```

Figure 1. Impact of Run Support on Wins

	(1)
	Wins in Ga~d
RS/GS	0.345***
	(0.102)
QS%	12.94***
	(1.630)
IP/GS	0.536***
	(0.137)
Constant	-3.021***
	(0.649)
Observations	370

Standard errors in parentheses

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

```
. esttab using regression_results.rtf, replace se label star(* 0.10 ** 0.05 *** 0.01)
(output written to regression_results.rtf)
```

.

.

. eststo simple_ols

.

. ***interaction model regression

. gen RSxQS = RSGS * U

. regress Wgs RSGS U IPGS RSxQS

```
      Source |       SS           df       MS      Number of obs   =       370
-------------+----------------------------------   F(4, 365)       =    151.69
       Model |  4780.07822         4  1195.01956   Prob > F        =    0.0000
    Residual |   2875.4461       365  7.87793452   R-squared       =    0.6244
-------------+----------------------------------   Adj R-squared   =    0.6203
       Total |  7655.52432       369  20.7466784   Root MSE        =    2.8068

------------------------------------------------------------------------------
         Wgs | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
        RSGS |   .0728993   .1019934     0.71   0.475    -.1276692    .2734678
           U |   2.970204   2.071541     1.43   0.152    -1.103449    7.043857
       IPGS |    .499202    .146419     3.41   0.001     .2112714    .7871327
      RSxQS |   2.264631   .4169455     5.43   0.000     1.444714    3.084547
      _cons |  -1.674432   .7103105    -2.36   0.019    -3.071247   -.2776173
```

. vif

Variable	VIF	1/VIF
U	9.78	0.102232
RSxQS	9.33	0.107189
IPGS	2.10	0.476597
RSGS	1.32	0.758547
Mean VIF	5.63	

.

. eststo inter_model_model

.

. ***log linear

. histogram Wgs, normal

(bin=19, start=0, width=.94736842)

. gen ln_Wgs = ln(Wgs + 1)

. regress ln_Wgs RSGS U IPGS, robust

Linear regression Number of obs = 370

F(3, 366) = 348.20

Prob > F = 0.0000

R-squared = 0.6764

Root MSE = .57104

| Robust

ln_Wgs | Coefficient std. err. t P>|t| [95% conf. interval]

---+---

RSGS | .0760114 .022843 3.33 0.001 .0310914 .1209315

U | 2.450047 .3305317 7.41 0.000 1.800067 3.100026

IPGS | .2282113 .0313233 7.29 0.000 .166615 .2898076

_cons | -.8039405 .1450291 -5.54 0.000 -1.089135 -.5187456

.

. eststo log_linear

. ***reduced model

. regress Wgs RSGS, robust

Linear regression Number of obs = 370

F(1, 368) = 7.07

Prob > F = 0.0082

R-squared = 0.0180

Root MSE = 4.5198

```
------------------------------------------------------------------------------
             |               Robust
         Wgs | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
        RSGS |   .3715591   .1397592     2.66   0.008     .0967323    .6463858
       _cons |   2.211479   .5789144     3.82   0.000     1.073084    3.349874
------------------------------------------------------------------------------
```

```
. eststo reduced_model

. ***report your results in a table

.

. esttab simple_ols reduced_model log_linear inter_model_model ///

>     using regression_comparison.rtf, replace ///

>     title("Comparison of Regression Results") ///

>     b(%9.3f) se label ///

>     star(* 0.10 ** 0.05 *** 0.01)

(output written to regression_comparison.rtf)

.

end of do-file

. save "C:\Users\Ibrahim.Kaba001\OneDrive - University of Massachusetts Boston\Econ452\final project\final projjj.dta"

file C:\Users\Ibrahim.Kaba001\OneDrive - University of Massachusetts Boston\Econ452\final project\final projjj.dta saved
```

. exit