Striking Gold in Software Repositories

An *Econometric* Study of Cryptocurrencies on GitHub

**Asher Trockman**, Rijnard van Tonder, Bogdan Vasilescu

MSR '19, May 26–27, Montréal, QC, Canada

Carnegie Mellon University
Why do we think there is “gold”? 

1.) Visible Signals on GitHub Influence Perceptions of Software Quality

Kristoufek 2013, 2015;

2.) “Attractiveness” Influences Crypto Speculators

Garcia et al. 2014

3.) Informed? Speculators Drive Price

Dabbish et al. 2012, Trockman et al. 2018
Collecting Crypto Financial & Development Data

...and 239 other cryptocurrencies:
(different availability for different parts of the study)

...and more!
Correlation with Avg. Market Cap = (price \times \text{coins})

**Popularity Metrics**

- Stars: 62%
- Watchers: 61%
- Forks: 65%

**Activity Metrics**

- Contrib.: 50%
- Commits: 43%
- LOC Added: 38%
- LOC Removed: 32%

---

**Quality Assurance Indicators**

- Any Badge: (0.22)
- Any CI: (0.30)

---

**Legend**

- Badge/CI: F
- True

---

**Y-Axis**

- Market Cap

**X-Axis**

- ...
## Linear Models of Avg. Market Cap

\[ \text{Avg. Market Cap} = (\text{price} \times \text{coins}) \]

### Popularity Metrics

<table>
<thead>
<tr>
<th>Metric</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stars</td>
<td>62%</td>
</tr>
<tr>
<td>Watchers</td>
<td>61%</td>
</tr>
<tr>
<td>Forks</td>
<td>65%</td>
</tr>
</tbody>
</table>

**Only popularity metrics are significantly and positively associated with avg. market cap.**

### Quality Assurance Indicators

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Coefficient</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu \log \text{Stars} )</td>
<td>0.29</td>
<td>0.18</td>
</tr>
<tr>
<td>( \mu \log \text{Watchers} )</td>
<td>0.03</td>
<td>0.20</td>
</tr>
<tr>
<td>( \mu \log \text{Forks} )</td>
<td>0.47</td>
<td>0.17</td>
</tr>
<tr>
<td>( \mu \log \text{Commits} )</td>
<td>0.35</td>
<td>0.30</td>
</tr>
<tr>
<td>( \mu \log + \text{LOC} )</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>( \mu \log - \text{LOC} )</td>
<td>0.08</td>
<td>0.18</td>
</tr>
<tr>
<td>( \mu \log \text{Contrib.} )</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>HasBadge</td>
<td>0.21</td>
<td>0.21</td>
</tr>
<tr>
<td>HasCI</td>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td>Intercept</td>
<td>14.0</td>
<td>0.48</td>
</tr>
</tbody>
</table>

**Joint***

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.29</td>
<td>0.18</td>
</tr>
<tr>
<td>0.03</td>
<td>0.20</td>
</tr>
<tr>
<td>0.47</td>
<td>0.17</td>
</tr>
<tr>
<td>0.35</td>
<td>0.30</td>
</tr>
<tr>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>0.08</td>
<td>0.18</td>
</tr>
<tr>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>0.21</td>
<td>0.21</td>
</tr>
<tr>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td>14.0</td>
<td>0.48</td>
</tr>
</tbody>
</table>

**Observations**: 149  
**Adj. \( R^2 \)**: 55.6%
Granger “Causality”

\[ y_t = \delta_0 + \sum_{j=1}^{p} \alpha_j y_{t-j} + \sum_{j=1}^{p} \gamma_j x_{t-j} + \epsilon_t \]

- **Dependent Variable**
- **Past Levels of Dependent Variable**
- **Past Levels of Independent Variable**

**Note:** *Toda-Yamamoto method*

**Coin: Bytom (1 of ~150)**

- Similar trends, for ~3 coins

**Past Levels of Dependent Variable:**
- **Market Cap**
- **Popularity**
- **Activity**

**Past Levels of Independent Variable:**
- **Stars**
- **Forks**
- **Watchers**

**Activity:**
- **Commits**
- **Added**
- **Removed**
We see evidence of Granger causality in only a few projects. Correcting for multiple hypotheses, this is insignificant.

**Stars** Granger-cause? Market Cap 9/142 coins
Binance, Cryptonex, Diamond, Electroneum, Emercoin, INS Ecosystem, Pandacoin, Vericoin, ZenCash

**Watchers** Granger-cause? Market Cap 4/146 coins
Ark, Mintcoin, NEM, PIVX

*(But we are likely to get similar from random noise.)*
Other models reveal a **very weak connection** between popularity and market cap, which is not robust.

**Stars** *Granger-cause? Market Cap*  9/142 coins

Binance, Cryptonex, Diamond, Electroneum, Emercoin, INS Ecosystem, Pandacoin, Vericoin, ZenCash

**Watchers** *Granger-cause? Market Cap*  4/146 coins

Ark, Mintcoin, NEM, PIVX

*(But we are likely to get similar from random noise.)*
Looks like we haven’t struck gold. BUT:

**Long-term: Signals of Pop. & QA**

<table>
<thead>
<tr>
<th>Correlation with Avg. Market Cap</th>
<th>(price × coins)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Popularity Metrics</strong></td>
<td></td>
</tr>
<tr>
<td>Stars</td>
<td>60%</td>
</tr>
<tr>
<td>Watchers</td>
<td>40%</td>
</tr>
<tr>
<td>Forks</td>
<td>20%</td>
</tr>
<tr>
<td><strong>Activity Metrics</strong></td>
<td></td>
</tr>
<tr>
<td>Contrib.</td>
<td>60%</td>
</tr>
<tr>
<td>Commits</td>
<td>40%</td>
</tr>
<tr>
<td>LOC Added</td>
<td>20%</td>
</tr>
<tr>
<td>LOC Removed</td>
<td>10%</td>
</tr>
</tbody>
</table>

**Quality Assurance Indicators**

- Any Badge (0.22)
- Any CI (0.35)

**Metrics ➞ Market Cap**

**Not compelling.**

*We see evidence of Granger causality in only a few projects.*

Correcting for multiple hypotheses, this is insignificant.

- **Stars Granger-cause? Market Cap**
  - 9/142 coins
  - Binance, Cryptonex, Diamond, Electroneum, Emercoin, INS Ecosystem, Pandacoin, Vericoín, ZenCash

- **Watchers Granger-cause? Market Cap**
  - 4/146 coins
  - Ark, Mintercoin, NEM, PIVX

*But we are likely to get similar from random noise.*

**Future work**

- Multi-year trends
- Sophisticated models
- Volatility or volume

Check out econometric techniques for future studies!

Contact
asher.trockman@gmail.com
http://ashertrockman.com

**Short-term: Very limited evidence**

**Activity** ➞ **Popularity** ➞ **Crypto Attractiveness** ➞ **Price**

**Quality Assurance**

**Representative example**

- **Monero**

**No Granger Causality**

**Build passing**