Меры для измерения скорости изменения размера фирм в задачах эконометрики

А. С. Семенова

Национальный исследовательский университет «Высшая школа экономики» — Санкт-Петербург

assemenova@hse.ru

Аннотация. В докладе обсуждаются различные меры, предназначенные для измерения скорости изменения размера фирм в задачах эконометрики. Выполнен их сравнительный корреляционный анализ на большом массиве данных, обсуждена показательная способность, выявлены особенности и очерчены границы областей возможного применения в эконометрических моделях. Исследована робастность соответствующих количественных значений рассмотренных мер к случайным погрешностям эконометрических данных.

Ключевые слова: размер фирмы, скорость изменения размера фирм, меры оценки роста предприятий, эконометрика

І. Введение

При решении задач эконометрики, связанных с анализом финансовых результатов фирм, ведущих предпринимательскую деятельность, одним из важных этапов является выбор переменных, релевантно и контрастно отображающих анализируемое измеримое свойство — будь то показатели полученной прибыли или характеристики кадрового состава компании. От выбора состава переменных зависит то, насколько использованная эконометрическая модель адекватна описываемым ею данным, нет ли неконтролируемых смещений в ее предсказаниях вследствие неполноты или зависимости переменных регрессионной модели.

Одними из важных переменных эконометрических моделей являются размер фирмы и метрики ее роста, поскольку они выступают детерминантами большого количества других эконометрических показателей. В литературе отмечается, что количество метрик для измерения размера фирм велико [1]. Действительно, основными показателями размера фирмы эмпирических эконометрических исследованиях выступает довольно большое количество различных величин: в первую очередь это количество сотрудников [2], общий объем продаж [3], реже – рост активов [4], отчетные значения продаж и прибыли [5]. Выводы обзора [1] говорят о том, что одной простой меры роста фирмы, к сожалению, предложить не удается. Кроме того, важными детерминантами роста выступают возраст фирмы [6, 7], объем ее инновационной деятельности [8, 9], прибыль [10], производительность [11], макро- и микроэкономические показатели [12-14]. Результаты работы [15] показывают, что выбранная метрика оценки роста фирмы должна определяться не только и не той переменной, которая детерминантой роста, но и длительностью промежутка времени, на которой данный рост требуется измерить.

К. К. Семенов

Санкт-Петербургский политехнический университет Петра Великого

semenov_kk@spbstu.ru

II. МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ МЕР РОСТА ФИРМ

Пусть G(t) — величина, характеризующая размер фирмы G в момент времени t. Тогда естественной метрикой изменения размера фирмы во времени выступает производная $V_0(t) = dG(t)/dt$. Поскольку время в эконометрике обычно откладывается в годах (из-за привязки к финансовой отчетности, формируемой лишь в определенные календарные периоды), то реально располагаемой для приложений метрикой может выступить конечная разность вида $V_1(t) = \Delta G(t)/\Delta t$, где $\Delta t = 1$ год, а $\Delta G(t)$ — изменение значения G за обозначенный период времени длительностью Δt . Как видим, на практике нет возможности оперировать со значениями $V_0(t)$, а лишь с их усредненными за период $\Delta t = 1$ год значениями $V_0(t)$:

$$\begin{split} V_1(t) &= \frac{\Delta G(t)}{\Delta t} = G(t + \Delta t) - G(t) = \int_t^{t + \Delta t} dG(t) = \\ \int_t^{t + \Delta t} V_0(t) dt &= \frac{1}{\Delta t} \cdot \int_t^{t + \Delta t} V_0(t) dt. \end{split}$$

Дискретность изменения переменной времени t сказывается на точности оценки: метрика $V_1(t)$ позволяет судить лишь о среднем увеличении или уменьшении размера фирмы, но не о наибольшей достигнутой скорости роста или падения. Для примерной оценки возможных отличий между $V_1(t)$ и значениями $\max_{\mathbf{\tau} \in [t,t+\Delta t]} V_0(\mathbf{\tau})$ и $\min_{\mathbf{\tau} \in [t,t+\Delta t]} V_0(\mathbf{\tau})$ максимальной и или минимальной скорости изменения за тот же период времени могут быть использованы следующие соображения. Поскольку согласно [1] в силу действия центральной предельной теоремы, значение G(t) как кумулятивная величина должно иметь распределение, близкое к нормальному, то и величина $V_1(t)$ должна распределение, близкое к нормальному. Следовательно, для оценки величины разности может быть использована статистика критерия Граббса [16] для регистрации статистических выбросов (outliers) в предположении нормального закона распределения. С доверительной вероятностью Q = 0.95 разности значений

$$\max_{\mathbf{\tau} \in [t,t+\Delta t]} V_0(\mathbf{\tau}) - V_1(t) \; \mathsf{i} \; V_1(t) - \min_{\mathbf{\tau} \in [t,t+\Delta t]} V_0(\mathbf{\tau})$$

не могут превысить из статистических причин значения

$$D = s \cdot \frac{N-1}{\sqrt{N}} \cdot \sqrt{\frac{t_{(1-Q)/(4\cdot N)}^2(N-2)}{N-2+t_{(1-Q)/(4\cdot N)}^2(N-2)}},$$
 (1)

где N < 248 — среднегодового количества рабочих дней, так как предполагается, что изменение размера фирмы не может фиксироваться чаще 1 раза в рабочий день

(например, предоставлением соответствующих сведений о штатной численности в надзорные ведомства); значение s есть среднеквадратическое отклонение ежедневных значений $V_0(t)$ в пределах анализируемого периода $\Delta t=1$ год; $t_{\beta}(v)-\beta\cdot 100$ %-ая квантиль распределения Стьюдента с числом степеней свободы, равном v.

Значение s может быть оценено по ежегодным значениям $V_1(t+k\cdot \Delta t)$, где $k=0,1,\ldots K-1$ следующим образом:

$$s^{2} = \frac{N}{K-1} \left(\sum_{k=0}^{K-1} V_{1}^{2}(t+k \cdot \Delta t) - K \cdot \sum_{k=0}^{K-1} V_{1}(t+k \cdot \Delta t) \right)$$

или, если данных недостаточно (K = 2), в предположении линейного роста (уменьшения) размера фирмы — как

$$s = s_0 \cdot |V_1(t + \Delta t) - V_1(t)|, \tag{2}$$

где $s_0 = 0,2899$ — значение среднеквадратического отклонения линейно растущей величины.

На значение оценки метрики изменения размера фирмы влияет не только дискретность значений переменной времени t, но и дискретность значений переменной размера G(t). Если в ее использована лишь качественная переменная, относящая фирму к категории малых, средних или больших (например, согласно [17, 18]), то мы имеем дело со следующей возникающей ошибкой. Пусть переход от малого размера к среднему характеризуется изменением величины G от G_0 до значения G_1 , а переход от среднего размера в большому – от G_1 до G_2 . Будем полагать, что уменьшение фирмы характеризуется переходами, но в обратную сторону: от G_2 к G_1 и от G_1 к G_0 . Тогда величина возникающей ошибки $\Delta V_1(t)$ равна

$$\Delta V_1^{i \to j}(t) = \frac{G_i - G_j}{\Delta t} - \Delta V_1(t) = \left(G_i - G_j - \Delta G(t)\right) / \Delta t, (3)$$

где индексы i и j отражают соответственно начальное и конечное значение размера фирмы на качественной шкале в анализируемом периоде времени.

Особенностью данной оценки является то, что требуемый перевод значений с качественной шкалы размеров фирмы «малый», «средний» и «большой» на количественную сам по себе сопряжен с неопределенностью и допущениями. При это линейность шкалы может и не соблюдаться, что приводит к дополнительным искажениям. В частности, при переходе от качественной шкалы к количественной может возникать явление гистерезиса (когда переход должен совершаться по-разному в зависимости от того, растет ли размер фирмы или, наоборот, уменьшается).

III. МЕРЫ ИЗМЕРЕНИЯ СКОРОСТИ РОСТА КОМПАНИЙ НА ОСНОВЕ ПОКАЗАТЕЛЕЙ С НИЗКОЙ ДИСКРЕТНОСТЬЮ

Основными проблемами, связанными с измерением скорости роста фирм, на основе количества работающих в ней сотрудников, являются сложность получения соответствующих сведений (далеко не всегда представленных в публичном доступе) и их усредненный характер (в частности использование среднесписочной численности персонала в бухгалтерской отчетности в качестве основной метрики). Это приводит к

возникновению тех проблем и затруднений, что были описаны в предыдущем разделе статьи.

Используем формулы (1) и (2) для демонстрации указанных обстоятельств. При росте численности сотрудников фирмы за один год от 10 до 50 человек (то есть от фирмы малого размера до средней фирмы) получаем по соотношению (2), что s = 11,6 человек. Как правило, в кадровой политике фирмы придерживаются политики планомерных изменений, не допускающих резких перепадов численности сотрудников. С известной долей общности можно допустить, что управленческие решения изменении штатного расписания об принимаются менеджментом компании примерно 1 раз в квартал, то есть N = 4. Из формулы (1) тогда следует, что величина возможных отклонений D, имеющих случайную природу, внутри года при этом с вероятностью Q = 0.95 не могла превосходить 13,4 человек, что представляет собой довольно значительную величину. Таким образом, неполнота и усредненный характер значений о численности сотрудников создают существенную статистическую неопределенность, мешающую успешно использовать в переменной размера фирмы G ее среднесписочную численность штатного расписания.

многих эконометрических базах содержащих сведения о финансовых результатах коммерческих фирм размер фирмы описывается в качественных градациях - малый (small), средний что также большой (large), (medium). затруднения - в частности приводит к возникновению гистерезиса при переходе к шкалам количественных значений. Пусть, например, используется, переход от качественной шкалы размера фирмы (малая, средняя, большая фирма) к значениям шкалы количества штатных сотрудников фирмы. Тогда при увеличении размера фирмы согласно [18] следует взять $G_1 = 50$ и $G_2 = 250$ в соответствии с наименьшей величиной штата сотрудников фирмы, позволяющей отнести ее к тому или иному размеру), а при убывании размера фирмы соответственно $G_0 = 10$ и $G_1 = 250$. При этом указанные значения могут отличаться от страны к стране в соответствии с действующим в них законодательством. Отмеченное обстоятельство присутствия гистерезиса должно учитываться эконометрическими моделями, однако этого не происходит, что не может не сказываться на результатах соответствующего анализа и выводах проводимых исследований.

IV. МЕРЫ ИЗМЕРЕНИЯ СКОРОСТИ РОСТА КОМПАНИЙ НА ДЛИТЕЛЬНЫХ ПЕРИОДАХ ВРЕМЕНИ

В случае, если в качестве метрики скорости роста фирмы используются финансовые показатели, то, как правило, объем данных для оценки соответствующих значений оказывается больше, чем в случае с количество сотрудников. Это позволяет строить более специфичные метрики для оценки значений, более полно учитывающие стохастический характер изменений фирмы при выполнении эконометрического анализа на большом объеме данных.

Обозначим финансовый показатель, используемый в качестве переменной роста фирмы, как f(t). Тогда среди метрик роста фирмы на протяжении заданного периода

 $[t, t+k\cdot \Delta t]$, где k>>1 могут быть использованы следующие:

- абсолютная или относительная разность значений f, взятых в первый и последний год в рассмотренной интервале времени: $\Delta_k f(t) = f(t+k\cdot\Delta t) f(t)$ и $\gamma_{f(t)} = \Delta f(t)/f(t)$;
- медианное значение ежегодных приращений f, высчитанное на анализируемом промежутке времени: Med = median $\{\Delta_1 f(\tau): \tau \in [t, t+(k-1)\cdot \Delta t]\};$
- разность между максимальным и минимальным значениями f, достигнутыми в изучаемом промежутке времени: Range = $\max_{\tau \in [t, t+k \cdot \Delta t]} f(\tau) \min_{\tau \in [t, t+k \cdot \Delta t]} f(\tau)$;
- коэффициент Slope наклона прямой, построенной с применением линейной регрессии методом наименьших квадратов, по значениям f в анализируемом временном промежутке.

Первая из перечисленных метрик эквивалентна усредненной оценки роста фирмы в обозначенном временном интервале и является наиболее широко применяемой на практике [1]. Ее значение позволяет понять, была ли фирма в среднем растущей в интервале $[t, t+k\cdot\Delta t]$ или нет: если $\Delta_k f(t)>0$, то это говорит о том, что на год $t+k\cdot\Delta t$ фирма выросла в сравнении с годом t; и наоборот – если $\Delta_k f(t)<0$, то это говорит о том, что на год $t+k\cdot\Delta t$ фирма уменьшилась в сравнении с годом t. Данная метрика является чувствительной к присутствию статистических выбросов в данных.

Вторая из представленных метрик более робастно отражает рост фирмы на рассматриваемом временном промежутке, поскольку использует ранговую статистику в своем составе. Значение метрики Med отражает главный тренд в изменении значений f во времени и не зависит от влияния внезапных макроэкономических явлений по типу «черного лебедя» [19]: если Med>0, то это указывает на то, что более чем в половине промежутка времени $[t, t+k\cdot\Delta t]$ фирма росла.

Разность между максимальным и минимальным значениями f в анализируемом интервале времени не является в достаточной мере робастной метрикой, но при этом она лучше отражает влияние на рост фирмы различных контролируемых факторов [15], включая инновационную деятельность. Известно [20], например, что наибольшее влияние от внедрения экологических инноваций на финансовые показатели фирмы достигается на второй год после внедрения — следовательно, значение Range должно лучше отражать влияние последних на значение f.

Наконец, четвертая из представленных метрик отражает главный (линейный) член зависимости значений f от времени. Ее значение позволяет отразить основной тренд в изменении размера фирмы: если Slope>0, то это указывает на то, что в целом фирма успешно росла на интервале времени $[t, t+k\cdot\Delta t]$. Данная метрика является обычным выбором при построении регрессионных эконометрических моделей.

Как видно, каждая из перечисленных метрик обладает своими преимуществами и недостатками и имеет разную чувствительность к неточностям в анализируемых данных (включая пропуски, нередко случающиеся в эконометрических базах данных). Это обстоятельство указывает на целесообразность применения сразу нескольких метрик при проведении соответствующих изысканий: совместное рассмотрение значений нескольких метрик позволит сделать более выводы при анализе на длительных глубокие промежутках времени по сравнению с использованием только одной из них. Несмотря на то, что значения перечисленных метрик вычисляются на одних и тех же данных, степень корреляции между ними довольно невысокая.

Для проверки этого утверждения были проведены специальные расчеты. Были использованы данные о показателях 3531 преимущественно европейских фирм за 2012–2019 гг. по базе данных Orbis (Bureau Van Dijk). описание использованных Подробное данных представлено в [20]. Были вычислены значения всех метрик, представленных в данном разделе для случаев, когда в качестве значения f были взяты чистая прибыль (net income) и общая стоимость активов (total assets). Для сравнения были также взяты значения качественной шкалы размера фирм по количеству ее сотрудников (small, medium, large). Была введена переменная S, чье значение для выросших в промежутке с 2012 по 2019 год фирм было равно (+1), для уменьшившихся в размере -(-1), для не изменившихся - соответственно 0. Были вычислены доверительные интервалы для значений коэффициента корреляции р между метриками при Q = 0.95. доверительной вероятности, равной Полученные результаты соответственно сведены в Таблицы 1 и 2.

ТАБЛИЦА I. ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ ДЛЯ КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ МЕЖДУ ЗНАЧЕНИЯМИ МЕТРИК РОСТА ФИРМ, ИЗМЕРЕННЫМИ ПО ЧИСТОЙ ПРИБЫЛИ

	$\Delta_8 f$	Med	Range	Slope	S
$\Delta_8 f$	[1,0; 1,0]	[0,61;0,65]	[-0,11;-0,05]	[0,90; 0,91]	[-0,06;0,00]
Med	[0,61;0,65]	[1,0; 1,0]	[-0,14;-0,07]	[0,60;0,64]	[-0,04;0,03]
Range	[-0,11;-0,05]	[-0,14;-0,07]	[1,0; 1,0]	[-0,11;-0,04]	[-0,03;0,04]
Slope	[0,90; 0,91]	[0,60;0,64]	[-0,11;-0,04]	[1,0; 1,0]	[-0,06;0,01]
S	[-0,06; 0,00]	[-0,04;0,03]	[-0,03;0,04]	[-0,06;0,01]	[1,0; 1,0]

ТАБЛИЦА II. ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ ДЛЯ КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ МЕЖДУ ЗНАЧЕНИЯМИ МЕТРИК РОСТА ФИРМ, ИЗМЕРЕННЫМИ ПО ОБЩЕЙ СТОИМОСТИ АКТИВОВ

	$\Delta_8 f$	Med	Range	Slope	S
$\Delta_8 f$	[1,0; 1,0]	[0,91;0,92]	[0,21;0,28]	[0,98;0,99]	[-0,03;0,03]
Med	[0,91;0,92]	[1,0; 1,0]	[0,16;0,22]	[0,90;0,91]	[-0,03;0,03]
Range	[0,21;0,28]	[0,16;0,22]	[1,0; 1,0]	[0,23;0,29]	[-0,02;0,04]
Slope	[0,98;0,99]	[0,90;0,91]	[0,23;0,29]	[1,0; 1,0]	[-0,03;0,04]
S	[-0,03;0,03]	[-0,03;0,03]	[-0,02;0,04]	[-0,03;0,04]	[1,0; 1,0]

В табл. 1 и 2 использована цветовая дифференциация, указывающая на силу корреляционной связи по шкале Чеддока: при $|\rho|<0,3$ — очень слабая связь (ярко-зеленый цвет); при $0,3<|\rho|<0,5$ — слабая связь; при $0,5<|\rho|<0,7$ — средняя связь (желтый цвет); при $0,7<|\rho|<0,9$ — высокая теснота связи; $|\rho|>0,9$ — очень сильная связь (красный цвет). Видим, что в целом результаты в таблицах повторяются: высокая корреляция между значениями

 $\Delta_8 f$ и Med говорит об отсутствии статистических выбросов в анализируемых данных, высокая корреляция между $\Delta_8 f$ и Slope – о присутствии выраженных трендов в изменении размера фирмы. Отсутствие корреляции между показателями роста, основанными на финансовых показателях фирмы, и на количестве сотрудников указывает на то, что данные показатели по-разному характеризуют размер фирмы, и на то, что по этой причине в эконометрических моделях нецелесообразно ограничиваться применением только одного из этих показателей.

Для оценки степени робастности приведенных в данном разделе метрик к случайным искажениям и погрешностям в анализируемых данных был также проведен следующий численный эксперимент. исходным значениям f был добавлен случайный шум с нормальным законом распределения с математическим ожиданием и варьируемым значением среднеквадратического отклонения (в пределах от 0 до 100% от среднего значения среднеквадратического отклонения значений f, взятых для каждой фирмы в отдельности, указанная доля обозначена как δ). Были построены односторонние доверительные интервалы для возникающей дополнительной среднеквадратической ошибки σ для доверительной вероятности Q = 0.95 в виде $[0, \sigma_{max}]$. Характер обозначенной зависимости является линейным, по этой причине результаты представлены в виде коэффициента порциональности между величиной σ_{max}/MAD , где MAD (median of absolute deviation from median) – медиана абсолютного отклонения OT медианы рассматриваемой метрики, вычисленная оригинальным исходным данным (т. е. без добавления случайного шума), и величиной б, представляющей собой, по сути дела, величину, обратную отношению «сигнал/шум» (SNR, signal-to-noise ratio). Полученные значения приведены в табл. 3 и характеризуют робастность рассмотренных метрик.

ТАБЛИЦА III. Коэффициент пропорциональности λ между значениями $\sigma_{\text{max}}/\text{MAD}$ и δ

	λ
$\Delta_8 f$	0.92
Med	0.15
Range	0.57
Slope	0.00

Как видим, наиболее робастной метрикой является Slope, чуть более чувствительны к искажениям значения Med, а наименее робастными являются, как и следовало ожидать, метрики Δf и Range.

V. ЗАКЛЮЧЕНИЕ

В данной работе исследованы особенности различных мер, используемых для измерения скорости изменения размера фирм в задачах эконометрики. Выполнен их сравнительный корреляционный анализ на большом массиве данных, обсуждена показательная способность, выявлены особенности и очерчены границы областей возможного применения в эконометрических моделях.

Исследована робастность соответствующих количественных значений рассмотренных мер к случайным погрешностям эконометрических данных.

Список литературы

- A. Coad and W. Hölzl, "Firm growth: Empirical analysis, Papers on Economics and Evolution," Report No. 1002, Max Planck Institute of Economics, Jena, 2010.
- [2] F. Delmar, "Measuring growth: Methodological considerations and empirical results," in Entrepreneurship and SME Research: On its Way to the Next Millennium (ed. by R. Donckels, and A. Miettinen), pp. 190–216, 1997.
- [3] K.M. Kiani, E.H. Chen and Z. Madjd-Sadjadi, "Financial factors and firm growth: evidence from financial data on Taiwanese firms," Quantitative Finance, vol. 12(8), pp. 1299-1314, 2012.
- [4] S.S. Cao, "Reexamining growth effects: are all types of asset growth the same?" Contemporary Accounting Research, vol. 33(4), pp. 1518-1548, 2016
- [5] R. Cressy, "Determinants of small firm survival and growth," in Oxford Handbook of Entrepreurship (ed. by M. Casson, B. Yeung, A. Basu, and N. Wadeson), Oxford Univ Press, pp. 161-193, 2006.
- [6] A. Coad and W. Holzl, "On the Autocorrelation of Growth Rates," Journal of Industry, Competition and Trade, vol. 9(2), pp. 139-166, 2009.
- [7] F. Lotti, E. Santarelli and M. Vivarelli, "Does Gibrat's Law hold among young, small firms?" Journal of Evolutionary Economics, vol. 13(3), pp. 213-235, 2003.
- [8] A. Coad and R. Rao, "Innovation and Firm Growth in High-Tech Sectors: A Quantile Regression Approach," Research Policy, vol. 37(4), pp. 633-648, 2008.
- [9] W. Holzl, "Is the R&D behaviour of fast-growing SMEs different? Evidence from CIS III data for 16 countries," Small Business Economics, vol. 33(1), pp. 59-75, 2009.
- [10] A. Moneta, D. Entner, P. Hoyer and A. Coad, "Causal inference by independent component analysis with applications to micro-and macroeconomic data", Report No. (2010, 031), Jena Economic Research Papers, 2010.
- [11] A. Pozzi and F. Schivardi, "Demand or productivity: What determines firm growth?" The RAND Journal of Economics, vol. 47(3), pp. 608-630, 2016.
- [12] R.J. Caballero and M. Hammour, "The Cleansing Effect of Recessions," American Economic Review, vol. 84, pp. 1350-1368, 1994.
- [13] D.C. Mead and C. Liedholm, "The Dynamics of Micro and Small Enterprises in Developing Countries," World Development, vol. 26, pp. 61-74, 1998.
- [14] D.B. Audretsch, "Innovation, growth and survival," International Journal of Industrial Organization, vol. 13, pp. 441-457, 1995.
- [15] A.S. Semenova, K.K. Semenov, M.A. Storchevoy, "Green Patents or Growth? Firms' Size Dynamics and Environmental Innovations Financial Gains," SSRN Electronic Journal, paper 10.2139/ssrn.4790916, 2024.
- [16] B.Y. Lemeshko and S.B. Lemeshko, "Extending the application of Grubbs-type tests in rejecting anomalous measurements," Measurement techniques, vol. 48(6), pp. 536-547, 2005.
- [17] OECD Enterprises by Business Size, 2024. URL: https://data.oecd.org/entrepreneur/employees-by-business-size.htm
- [18] "What is an SME? Small and medium sized enterprises (SME) Enterprise and Industry". URL: ec.europa.eu.
- [19] X. Liu, "The Economic Environment and Black Swan Events: Taking COVID-19 as an example," Journal of Education, Humanities and Social Sciences, vol. 23, pp. 336-341, 2023.
- [20] A.S. Semenova, K.K. Semenov and M.A. Storchevoy, "One, Two, Three: How Many Green Patents Start Bringing Financial Benefits for Small, Medium and Large Firms?" Economies, vol. 11(5), paper 137, 2023