

М. СТОЛБОВ,  
кандидат экономических наук,  
доцент кафедры прикладной экономики МГИМО

## СТАТИСТИКА ПОИСКА В GOOGLE КАК ИНДИКАТОР ФИНАНСОВОЙ КОНЪЮНКТУРЫ

Глобальный кризис 2008—2009 гг. поставил перед экономической наукой ряд сложных теоретико-методологических и прикладных задач. Повышение качества мониторинга и прогнозирования деловой активности представляет собой одно из направлений исследований, где эти задачи тесно смыкаются. Повышение точности прогнозов и увеличение упреждающих интервалов для принятия решений важны для успеха стабилизационной (контрциклической) политики государства.

Традиционно для заблаговременного выявления экономической нестабильности используют системы опережающих индикаторов, или сводные опережающие индикаторы (*composite leading indicators*). Разработка подобных показателей представляет собой динамично развивающуюся, но уже прошедшую путь институционализации область прикладных экономических исследований<sup>1</sup>. Большинство сводных опережающих индикаторов нацелено на отслеживание *общеекономической* конъюнктуры, поэтому при их построении тенденции развития отдельных секторов экономики могут накладываться друг на друга. Для получения более обоснованных оценок целесообразно использовать несколько альтернативных сводных опережающих индикаторов, а также аналогичные «барометры», характеризующие состояние отдельных секторов экономики.

В последние два десятилетия из секторальных барометров все большую популярность приобретают *индикаторы финансовой конъюнктуры* (FCI, *financial conditions indices*). На уровне национальных денежных властей такие показатели применяются в США, Канаде, Швейцарии. Эти обобщающие индикаторы основаны на применении факторного анализа, который проводится на базе 30—40 исходных показателей, учитывающих конъюнктуру рынка ценных бумаг, финансовых посредников, рынка недвижимости, влияние мер денежно-кредитной политики<sup>2</sup>. Наиболее успешные индикаторы финансовой конъюнктуры обладают упреждающим потенциалом, равным двум-трем кварталам<sup>3</sup>.

Вместе с тем следует учитывать, что фактический горизонт упреждения оказывается все же меньше: при построении индикаторов

---

<sup>1</sup> Деятельность экспертов в построении опережающих индикаторов интегрируется Центром международных конъюнктурных исследований и опросов (CIRET, Centre for International Research on Economic Tendency Surveys), поддерживается статистическими службами ООН и ОЭСР. Кроме того, данное профессиональное сообщество имеет научный журнал — *Journal of Business Cycle Measurement and Analysis*.

<sup>2</sup> *English W., Tsatsaronis K., Zoli E.* Assessing the Predictive Power of Measures of Financial Conditions for Macroeconomic Variables // BIS Paper. 2005. No 22. P. 243—248.

<sup>3</sup> *Hatzius J., Hooper P., Mishkin F., Schoenholtz K., Watson M.* Financial Conditions Indexes: A Fresh Look After the Financial Crisis // NBER Working Paper. 2010. No 16150.

финансовой конъюнктуры используется официальная статистическая информация, сбор, обработка и публикация которой сопряжены со значительными временными лагами. Данное обстоятельство снижает эффективность этих показателей, причем кардинально в случае стран с недостаточно развитой системой государственной статистики и надзора за финансовой системой. Преодолеть эти сложности можно, используя высокочастотные данные (как правило, биржевые показатели) и результаты социологических опросов (разнообразные индексы финансовых настроений и т. п.).

В данной работе мы предлагаем альтернативу — анализ динамики специфических запросов интернет-пользователей в поисковой системе Google как инструмент *выявления* финансовых настроений экономических агентов. Запросы можно использовать для оценки текущей финансовой конъюнктуры и краткосрочного прогнозирования будущей.

### **Запросы в системе Google как эффективный инструмент текущей оценки и прогноза рыночной ситуации**

Через поисковую систему Google осуществляется около 70% всех интернет-запросов в мире. При этом данная компания внедрила мощный и публично доступный инструмент анализа статистики пользовательских запросов — Google Insights ([www.google.com/insights/search](http://www.google.com/insights/search)). С его помощью можно получить доступ к статистике запросов по широкому перечню стран, начиная с 2004 г. Пользовательские запросы объединены в тематические категории и подкатегории. Например, запросы, так или иначе относящиеся к финансам, сведены в категорию «Финансы и страхование», включающую семь подкатегорий («Банковское дело и личные вклады», «Бухгалтерский учет и налоги», «Денежные единицы и иностранная валюта», «Инвестиции», «Кредиты и займы», «Пенсия», «Страхование»). Особенность Google Insights состоит в возможности визуализации динамики запросов по определенному дескриптору (или группе дескрипторов) и экспорта данных в формате CSV<sup>4</sup>.

В 2009 г. Х. Чой и Х. Вэриан выдвинули гипотезу о том, что статистика запросов в Google должна коррелировать с *текущим* уровнем деловой активности, а также может быть полезной для его краткосрочного прогноза. Как отмечается в их исследовании, количество запросов в отношении определенной марки автомобилей за вторую неделю июня может помочь в прогнозировании объема продаж за июнь в целом<sup>5</sup>. Они обосновали свое предположение на примере продаж автомобилей Ford, Chevrolet и Toyota на рынке США, динамики туристического потока в Гонконг и ряда других показателей.

---

<sup>4</sup> Google Insights не выдает данные в формате абсолютных значений частоты запросов. Вместо них используются нормализованные относительно исходного периода (января 2004 г.) показатели. Максимальное значение нормализованной частоты запросов равно 100, минимальное — 0. Подробнее о статистической методологии Google Insights см.: [www.google.com/insights/search/#](http://www.google.com/insights/search/#).

<sup>5</sup> Choi H., Varian H. Predicting the Present with Google trends. Mimeo. April 2009.

Методология анализа была довольно проста: сначала строились авторегрессионные модели для каждой переменной, затем в эти модели добавляли динамику запросов в Google по соответствующему дескриптору как независимую переменную и сопоставляли качество исходной и усовершенствованной моделей по критерию скорректированного коэффициента детерминации и величине средней абсолютной ошибки, выраженной в процентах (MAPE, mean average percent error). Во всех числовых примерах, приведенных в работе Чоя и Вэриана, модели с включением переменной, отвечающей за динамику запросов в Google, превосходили исходные авторегрессии по указанным критериям.

Пионерное исследование американских экономистов вызвало интерес в исследовательском сообществе. С момента его публикации до июня 2011 г. в систему исследовательской сети по социальным наукам (SSRN) были введены данные о 16 работах, выполненных «по мотивам» исследования Чоя и Вэриана и содержащих соответствующую ссылку. Эти работы можно разделить на три группы: исследования, тестирующие возможности запросов в Google с точки зрения маркетинга; макроэкономические исследования; эмпирические исследования по финансовой экономике и корпоративным финансам. Осознавая значимость и перспективность учета динамики запросов в Google в маркетинговых исследованиях, мы уделим основное внимание второй и третьей группам.

В макроэкономических исследованиях, пожалуй, главная сюжетная линия связана с использованием динамики запросов в Google с целью прогнозировать потребление домохозяйств. Такие исследования были проведены для США, Германии и Израиля. На основе американских данных был сделан вывод о предпочтительности прогнозных моделей, учитывающих динамику запросов в Google, по отношению к прочим модельным постановкам. Более того, в периоды экономической нестабильности (в 2008–2009 гг.) их преимущество было наиболее ощутимым, в том числе по сравнению с моделями, где в качестве предиктора потребления домохозяйств использована динамика таких известных индексов потребительской уверенности, как индекс потребительских настроений Мичиганского университета<sup>6</sup>.

Особенностью исследования Т. Шмидта и С. Возена по Германии стала разработка с помощью факторного анализа индикатора потребительских настроений, построенного *исключительно* на данных запросов в Google по 46 категориям поиска, имеющим отношение к потребительским расходам<sup>7</sup>. Как в случае США, построенный индекс сопоставлялся с ведущими европейскими индексами потребительских настроений — индикатором потребительской уверенности Европейской комиссии и индикатором уверенности в розничной торговле — и доказал свое превосходство.

---

<sup>6</sup> Kholodilin K., Podstawski M., Siliverstov B., Burgi C. Google Searches as a Means of Nowcasts of Key Macroeconomic Variables // DIW Berlin Discussion Paper. 2009. No 46; Kholodilin K., Podstawski M., Siliverstov B. Do Google Searches Help in Nowcasting Private Consumption? A Real-Time Evidence for the US // KOF Swiss Economic Institute Working Paper. 2010. No 256.

<sup>7</sup> Schmidt T., Vosen S. A Monthly Consumption Indicator for Germany Based on Internet Search Query Data // Ruhr Economic Paper. 2010. No 208.

Помимо потребления домохозяйств, с помощью динамики запросов в Google были построены прогнозы уровней безработицы в США, Германии, Израиле и количества обращений за пособиями по безработице в этих странах<sup>8</sup>. Изучение динамики обращений за пособиями по безработице с помощью запросов в Google — пример перехода к макроэкономическому мониторингу и прогнозированию на еженедельных данных.

Использование статистики пользовательских запросов в Google открывает новые возможности в финансовой экономике и корпоративных финансах. При построении различных моделей ценообразования активов (прежде всего, ценных бумаг) возникает проблема подбора переменных, адекватно фиксирующих сдвиги в *предпочтениях* и *внимании* инвесторов. В качестве инструментальных переменных, которые отражают внимание инвесторов, предлагались экстремальные значения доходности активов, объем торгов по ним, расходы на рекламу, связанную с торгуемым активом, частота упоминаний о компании/активе в СМИ и т. п. Но перечисленные показатели косвенно измеряют внимание инвесторов.

Интенсивность поисковых запросов в Google в отношении определенного актива, напротив, непосредственно измеряет внимание инвесторов при условии достаточной ликвидности рынка этого актива, множественности участников и относительно симметрично распределенной информации. Перечисленные условия фактически описывают, насколько рынок отвечает гипотезе эффективности по Ю. Фаме. Американский фондовый рынок со множеством частных инвесторов, которые регулярно ищут информацию в Интернете в отношении финансовых инструментов, входящих в их портфели, тяготеет к сильной форме эффективности. Таким образом, динамика запросов в Google, вероятно, способна фиксировать изменение внимания инвесторов к тому или иному активу.

Данная гипотеза нашла эмпирическое подтверждение. Анализ взаимосвязи динамики котировок акций 3000 публичных компаний в США и запросов по ним в Google показал, что рост количества последних приводит к повышению курсовой стоимости акций компаний в течение двух недель с последующим возвращением к исходному ценовому уровню в пределах одного года<sup>9</sup>. Также было установлено, что повышенное число запросов в Google, с одной стороны, ассоциируется с резким ростом курса акций в первый день торгов после проведения IPO, а с другой стороны, сочетается с уровнем цен ниже цены размещения в долгосрочном периоде. Кроме того, отмечается, что динамика интернет-запросов положительно коррелирует с перечисленными выше косвенными показателями внимания инвесторов. Наиболее выражена взаимосвязь с объемом

---

<sup>8</sup> Choi H., Varian H. Predicting Initial Claims for Unemployment Benefits. Mimeo. 2009. April; Suhoy T. Query Indices and a 2008 Downturn: Israeli Data // Bank of Israel Technical Report. 2009. April; Askitas N., Zimmerman K. Google Econometrics and Unemployment Forecasting // Applied Economics Quarterly. 2009. Vol. 55, No 2. P. 107–120.

<sup>9</sup> Da Z., Engelberg J., Gao P. In Search of Attention // Journal of Finance. 2011. Vol. 66, No 5. P. 1461–1499.

торгов<sup>10</sup>. Но при совместном включении в регрессионные модели динамика интернет-запросов лучше предсказывает конъюнктуру, чем косвенные показатели.

Было показано, что анализ запросов в Google чрезвычайно актуален для исследований в корпоративных финансах<sup>11</sup>. Поскольку интенсивность поиска определенного бренда (или группы брендов) в Интернете помогает оценивать их фактический объем продаж, она в значительной степени определяет и операционный денежный поток компании-производителя. Сформировав выборку из 865 публичных компаний, исследователи определили «главный» продукт для каждой из них, исходя из интенсивности рекламы. Затем для каждого продукта была собрана статистика запросов в Google за 2004–2008 гг. Динамика интернет-запросов оказалась хорошим предиктором выручки, особенно для монопродуктовых компаний и фирм с ограниченным количеством брендов, а также динамично растущих фирм («газелей»). Так как выручка и производные от нее бухгалтерские показатели играют важную роль в моделях инвестиционной оценки, использование интернет-запросов может существенно повысить их прогнозный потенциал.

Перспективным представляется привлечение статистики запросов в Google для оценки макрофинансовой устойчивости национальных экономик. Повышению качества так называемого макропруденциального надзора уделяется существенное внимание в деятельности G-20, МВФ, Совета по финансовой стабильности и других международных экономических организаций. Одна из главных трудностей, как и в случае со сводными опережающими индикаторами общеэкономической конъюнктуры, заключается в значительном сужении упреждающего горизонта в связи с лагами выхода финансовых данных. Попытаемся использовать статистику запросов в Google для оценки финансовой конъюнктуры в российской экономике.

### **Применение статистики запросов в Google для оценки финансовой конъюнктуры в России**

Соответствующий анализ проведем по трем направлениям. Сначала докажем целесообразность использования статистики запросов в Google<sup>12</sup> для объяснения динамики отдельных агрегированных финансовых показателей (на примере депозитов физических лиц). Затем сконструируем «барометр» финансовой конъюнктуры для отечественной экономики, который основан исключительно на запросах

---

<sup>10</sup> *Preis T., Reith D., Stanley E.* Complex Dynamics of our Economic Life on Different Scales: Insights from Search Engine Query Data // *Philosophical Transactions of the Royal Society*. 2010. Vol. 368, No 1933. P. 5707–5719.

<sup>11</sup> *Da Z., Engelberg J., Gao P.* In Search of Fundamentals. Mimeo. 2011. March.

<sup>12</sup> Необходимо отметить, что в отличие от международного рынка поисковых систем, в России Google занимает вторую позицию после Яндекса, имея рыночную долю примерно 25%. Яндекс также внедрил систему статистической обработки запросов пользователей, однако их продолжительность по сравнению с Google невелика — с июня 2009 г. На наш взгляд, статистика запросов в Google достаточна репрезентативна, хотя из-за различий в методологии ее невозможно сопоставить с результатами Яндекса.

интернет-пользователей по специфическим дескрипторам, отражающим изменчивость их финансовых настроений. Также проверим, как предложенный сводный индикатор соотносится с индикаторами финансовых настроений, базирующимися на данных социологических опросов (на примере Индекса финансовых настроений Сбербанка РФ).

*Запросы в Google и объем привлеченных банками вкладов (депозитов) физических лиц*

На 1 июля 2011 г. доля вкладов физических лиц в структуре пассивов банковской системы составляла 30%<sup>13</sup>. Данная статья пассивов является ключевой для фондирования кредитной активности большинства российских банков. Важно отметить, что физические лица выступают нетто-кредиторами банковской системы: объем выданных им ссуд почти вдвое меньше размещенных ими вкладов. Таким образом, привлеченные средства физических лиц обеспечивают кредитование юридических лиц и иные банковские операции. Их динамика имеет существенное значение для поддержания устойчивости банковской системы и макрофинансовой устойчивости российской экономики.

Ниже приведены значения абсолютного прироста объемов вкладов физических лиц в банковской системе за январь 2004 — июнь 2011 г. (рис. 1). На графике четко показаны влияние сезонного фактора на динамику привлечения депозитов, а также повышательный тренд. Кроме того, есть основания предполагать, что модель привлечения депозитов в начальный период кризиса 2008—2009 гг. претерпела некий структурный сдвиг.



Рис. 1

Сезонность и тренд наблюдаются и в интенсивности запросов в Google, за исключением периода с января 2004 по январь 2006 г., когда их нормализованная частота была почти равна нулю (рис. 2).

<sup>13</sup> Рассчитано по данным Обзора банковского сектора Российской Федерации No 106 (август 2011 г.). [cbr.ru/analytics/bank\\_system/obs\\_1108.pdf](http://cbr.ru/analytics/bank_system/obs_1108.pdf).



Рис. 2

Для подтверждения обоснованности построения МНК-регрессии ежемесячного прироста депозитов на интенсивность запросов в Google за первую неделю каждого месяца необходимо, чтобы оба этих временных ряда относились к так называемым стационарным процессам с детерминированным трендом (trend stationary processes). Такой результат позволил бы исключить вероятность подмены реальной экономической взаимосвязи между переменными «кажущейся» регрессией<sup>14</sup>.

С инструментальной точки зрения, для выявления нестационарных процессов с детерминированным трендом используются тесты на наличие единичного корня. Воспользуемся наиболее популярным из них — расширенным тестом Дики—Фуллера. Он позволяет уверенно (с вероятностью выше 99%) отвергнуть гипотезу о наличии единичного корня в случае обоих временных рядов с учетом константы и временного тренда. Таким образом, взаимосвязь между абсолютным приростом депозитов физических лиц и интенсивностью запросов пользователей в Google правомерно моделировать с помощью метода наименьших квадратов. Поэтому методику Вэриана—Чоя в нашем случае можно применять.

Наиболее адекватно учесть сезонность в динамике депозитов физических лиц позволяет авторегрессионная модель AR(12), уравнение которой за весь период выглядит как  $DEP = 48\,471,79 + 0,69DEP(-12)$ <sup>15</sup>. Масштабный отток вкладов, зафиксированный в октябре 2008 г. (354,5 млрд руб.), является «выбросом», который статистически можно передать с помощью фиктивной переменной. Это позволяет заметно повысить качество модели. Регрессионное уравнение теперь выглядит так:  $DEP = 55\,930,49 + 0,67DEP(-12) - 44\,4154,9DUMMY$ <sup>16</sup>.

<sup>14</sup> Канторович Г. Г. Анализ временных рядов. Лекция 7 // Экономический журнал ВШЭ. 2002. Т. 6, № 2. С. 268—273.

<sup>15</sup> Скорректированный коэффициент детерминации составил 0,35. Значение *t*-статистики свободного члена равно 2,2, коэффициента при лагированном абсолютном приросте депозитов — 3,4. *F*-статистика уравнения значима на уровне 1%.

<sup>16</sup> Скорректированный коэффициент детерминации составил 0,52. Значение *t*-статистики свободного члена равно 2,7, коэффициента при лагированном абсолютном приросте депозитов — 3,4, при фиктивной переменной — 35,3. *F*-статистика значима на уровне 1%.

Для обеих регрессий тест Харке—Бера отклоняет гипотезу о нормальном распределении остатков, по крайней мере, на уровне значимости 5%. Поэтому качество полученных зависимостей в дальнейшем можно улучшить.

Проверим с помощью теста Чоу, насколько правильно строить единую модель для всего периода наблюдений. График подсказывает, что возможный структурный сдвиг связан с резким оттоком депозитов физических лиц в октябре 2008 г. Применение данного теста подтверждает это предположение. Целесообразно оценивать две модели — за январь 2004—октябрь 2008 г. и ноябрь 2008—май 2011 г.

Для первого периода регрессионное уравнение записывается как  $DEP = 12\,386,14 + 1,07DEP(-12) - 421\,203,9DUMMY$ . Константа в модели утрачивает статистическую значимость, но общее качество модели заметно повышается: скорректированный коэффициент детерминации возрос до 0,79,  $t$ -статистики составили 8,4 и -67,7.  $F$ -статистика уравнения значима на уровне 1%.

Во втором периоде статистически значимое влияние на привлечение депозитов в текущем месяце оказывал показатель с лагом десять месяцев:  $DEP = 55\,060 + 0,29DEP(-10) + 0,58DEP(-12)$ . Скорректированный коэффициент детерминации равен 0,39,  $t$ -статистики составили 1,8, 2,1 и 4,4 соответственно.  $F$ -статистика уравнения значима на уровне 1%. Для обеих регрессий теперь остатки, согласно тесту Харке—Бера, распределены нормально.

Проверим, насколько статистика запросов по дескриптору «вклады (депозиты)»<sup>17</sup> поможет уточнить полученные регрессионные модели. По аналогии с исследованием Чоя и Вэриана, мы предполагаем наличие лага между поиском информации в Google и фактическим открытием банковского вклада. Поэтому в качестве независимой переменной использована интенсивность запросов за первую неделю отчетного месяца<sup>18</sup>.

В рамках первого периода динамика запросов в Google не является статистически значимой переменной. Уравнение принимает вид  $DEP = 16\,620,6 + 1,11DEP(-12) - 416\,259,2DUMMY - 204,04GOOGLE$ . Скорректированный коэффициент детерминации остался на уровне 0,79,  $t$ -статистики равны 1,9, 8, -46 и -1 соответственно. Кроме того, коэффициент при динамике запросов в Google имеет «неправильный» знак. Такой результат обусловлен тем, что примерно до начала 2006 г. интенсивность запросов по дескриптору «вклады (депозиты)» в Google среди российских пользователей была близка к нулю. В свою очередь, это, вероятно, можно связать с еще не укоренившейся на тот момент традицией поиска финансовой информации и недостаточной активностью российских банков по продвижению своих услуг в Интернете. В частности, по данным Росстата, число пользователей

<sup>17</sup> Наряду с этим дескриптором, Google Insights определяет как популярные запросы по дескрипторам «банк вклады», «банковские вклады» и «вклад». Коэффициенты корреляции запросов по использованному нами дескриптору и по приведенным аналогам находятся в пределах 0,85—0,98. Таким образом, вряд ли тонкости формулировки оказывают в нашем случае существенное воздействие на результат (ключевое слово «вклад»).

<sup>18</sup> Альтернативной эконометрической методологией здесь выступает аппарат MIDAS-регрессий, предложенный в работах Э. Гиселса, быстро набирающий популярность в финансовой эконометрике и прогнозировании. См., например: Ghysels E., Sinko A., Valkanov R. MIDAS Regressions: Further Results and New Directions // *Econometric Reviews*. 2007. Vol. 26, No 1. P. 53—90. Для расчетов в данной статье использован пакет EViews 5.1, который, к сожалению, не располагает автоматизированной возможностью реализовать данную процедуру.

на 1000 населения возросло за 2000–2005 гг. с 20 до 153 человек. С точки зрения теории сетевых эффектов в 2005–2006 гг. сформировалась критическая масса пользователей, обусловившая динамичный рост в дальнейшем (рис. 3). Значительные изменения за этот период произошли в использовании интернет-технологий кредитными организациями: к концу 2005 г. количество банков, установивших системы дистанционного банковского обслуживания через Интернет, превысило 40% общего числа зарегистрированных кредитных организаций и почти удвоилось за 2003–2005 гг.<sup>19</sup>



Источник: данные Росстата, Минкомсвязи.

Рис. 3

Поэтому объяснимо, что во втором временном периоде нормализованная частота поиска дескриптора «вклады (депозиты)» в Google помогает объяснить абсолютный прирост вкладов физических лиц. Уравнение регрессии имеет вид:  $DEP = -98\ 002 + 0,21DEP(-10) + 0,53DEP(-12) - 2758GOOGLE$ . В результате включения динамики запросов в Google в модель скорректированный коэффициент детерминации увеличился с 0,39 до 0,47. Значения *t*-статистик составили 1,0, 2,1, 2,7 и 2,3. Сопоставляется с интуитивными ожиданиями и знак коэффициента при динамике запросов: их интенсивность должна стимулировать приток депозитов в банковскую систему по итогам месяца.

Построенные модели вряд ли можно эффективно использовать для прогнозных целей вне временных интервалов, по которым они построены. Значения коэффициентов детерминации, характеризующие долю объясненной вариации динамики роста депозитов, все же далеки от 1. После всех улучшений остается заметной и средняя абсолютная ошибка регрессии (47%). Однако расчеты внутри выборки (*in-sample*) указывают на значимую роль интернет-запросов в оценке динамики депозитов. Судя по уравнению регрессии за ноябрь 2008 — июнь 2011 г., увеличение нормализованной частоты запросов в Google за первую неделю отчетного месяца в среднем «добавляет» почти 2,8 млрд руб. прироста объема депозитов по его итогам. Поэтому можно констатировать целесообразность учета интенсивности запросов в Google при моделировании привлечения вкладов физических лиц банковской системой России.

<sup>19</sup> Данные портала CNews.ru.

*Построение барометра финансовой конъюнктуры  
с помощью Google и его соотношение с Индексом  
финансовых настроений Сбербанка РФ*

Интенсивность поиска информации в Google, очевидно, может не только коррелировать с отдельными финансовыми агрегатами, но и отражать степень уверенности экономических агентов в устойчивости финансовой системы в целом. Для подобной «диагностики» необходимо обобщить динамику запросов пользователей по наиболее значимым дескрипторам, способным подавать адекватные сигналы о состоянии финансовой конъюнктуры. Иными словами, нужно построить интегральный показатель финансовой конъюнктуры на базе запросов по индивидуальным дескрипторам.

Как определить перечень индивидуальных дескрипторов? Предлагается в качестве основы принять перечень дескрипторов из категории «Финансы и страхование», в отношении которых наиболее интенсивно осуществлялся поиск в острый кризисный период, то есть с сентября 2008 по июнь 2009 г. Глобальный кризис негативно повлиял практически на все сегменты российской финансовой системы. Поэтому настроения экономических агентов, преобладавшие в тот период и нашедшие выражение в интенсивности поиска по определенным дескрипторам в Google, на наш взгляд, можно экстраполировать на последующие эпизоды финансовой нестабильности.

На основе предложенного критерия были отобраны дескрипторы: «банк», «курс доллара», «евро», «ЦБ», «РТС», «ММВБ». Он был дополнен за счет дескрипторов — лидеров по запросам внутри отдельных подкатегорий категории «Финансы и страхование», не вошедших в наиболее востребованные в целом по этой категории, но важных для характеристики конъюнктуры определенного сегмента финансовой системы. В результате исходный список пополнился такими дескрипторами, как «акции», «ипотека», «взять кредит», «ПИФ», «банкротство», «залог», «дефолт», «девальвация», «финансовый кризис». Следует особо подчеркнуть, что составление перечня индивидуальных дескрипторов удалось формализовать благодаря встроенному в Google Insights механизму выявления десяти самых востребованных дескрипторов в рамках категорий и подкатегорий за определенный период.

Интегральный показатель состояния (барометр) финансовой конъюнктуры представляет линейную комбинацию запросов по указанным 15 индивидуальным дескрипторам с определенными весами:

$$Y = \omega_1 x_1 + \omega_2 x_2 + \dots + \omega_{15} x_{15},$$

где:  $Y$  — значение барометра финансовой конъюнктуры,  $x_i$  — нормализованная динамика запросов в Google по  $i$ -му дескриптору в среднем за месяц,  $\omega_i$  — вес динамики запросов в Google по  $i$ -му дескриптору в интегральном показателе.

Веса определены по формуле:

$$\omega_i = \sum_{i=1}^{14} r_{ij} / \sum_{i=1}^{14} \sum_{j=1}^{14} r_{ij},$$

то есть сумма коэффициентов парной корреляции каждого показателя и остальных соотносилась с общей суммой коэффициентов по матрице коэффициентов парной корреляции. Поскольку последняя отражает взаимосвязь между всеми показателями, полученные величины отражают удельный вес каждого показателя в общей величине интегрального индикатора. Веса рассчитаны по данным за январь 2004—январь 2011 г. (табл. 1)

Т а б л и ц а 1

**Веса динамики запросов в Google по индивидуальным дескрипторам в барометре финансовой конъюнктуры**

Дескриптор	W
Евро	0,079
Курс доллара	0,045
Банк	0,049
ЦБ	0,016
Взять кредит	0,004
Ипотека	0,069
Акции	0,035
РТС	0,110
ММВБ	0,120
ПИФ	0,044
Залог	0,089
Дефолт	0,101
Девальвация	0,067
Финансовый кризис	0,078
Банкротство	0,094

С учетом однородности значений индивидуальных показателей видно, что при одинаковом изменении запросов в Google наибольшее влияние на изменение барометра финансовой конъюнктуры оказывают настроения на *фондовом рынке* (дескрипторы «РТС» и «ММВБ») и *общеекономические* ожидания, выраженные через дескрипторы «дефолт» и «банкротство». Взвешивая значения показателей за каждый месяц и суммируя их, получаем величину барометра финансовой конъюнктуры.

Сконструированный барометр позволяет довольно уверенно выявлять *масштаб* текущих изменений в финансовой системе России. В частности, отчетливо заметны всплеск интереса пользователей Интернета к финансовым рынкам и инструментам в первой половине 2006 г. (когда запросы по дескрипторам «ПИФ», «акции», «ипотека» многократно возросли), период «спокойствия», длившийся до августа 2008 г., и резкий рост значений барометра в связи с финансовым кризисом (рис. 4). График позволяет определить временные границы его острой фазы — с сентября 2008 г. по апрель 2009 г. Кроме того, видно, что общей тенденцией в течение 2010—начале 2011 г. стало снижение поисковой активности интернет-пользователей в отношении финансовых дескрипторов. Это следует расценивать, с одной стороны, как отсутствие выраженных кризисных ожиданий, а с другой — как недостаточное доверие к финансовым рынкам и инструментам. Можно это воспринимать как «шлейф» кризиса. Напряженная ситуация вокруг лимита внешнего долга США в конце июля 2011 г. привела по ито-

**Динамика барометра финансовой конъюнктуры  
в России, январь 2004 — июль 2011 г.**

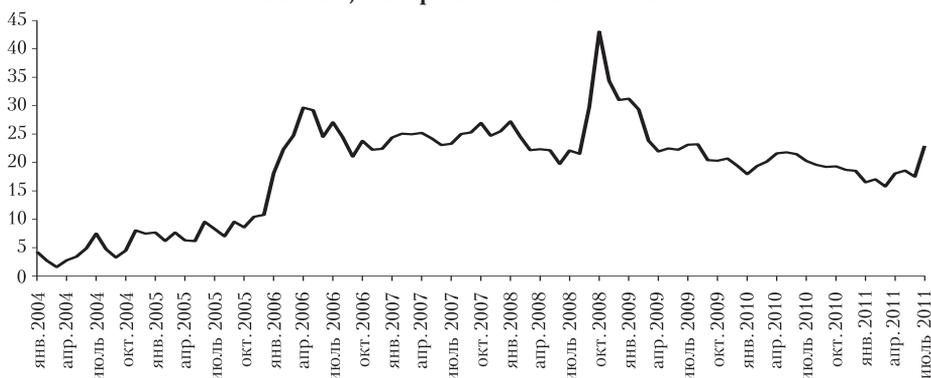


Рис. 4

гам месяца к возвращению барометра на уровень лета 2009 г. Резкий рост барометра финансовой конъюнктуры был связан с многократным увеличением поиска по дескрипторам «дефолт», «банкротство» и «девальвация». Динамика поиска по другим ключевым словам в целом не нарушила ранее сложившиеся тенденции<sup>20</sup>.

Альтернатива нашему подходу — метод главных компонент. Впрочем, полученные на его основе выводы не противоречат результатам применения нашей методики. Коэффициент корреляции первой главной компоненты, полученной на основе избранных 15 дескрипторов, и барометра финансовой конъюнктуры, составляет 0,97. Как и в случае барометра финансовой конъюнктуры, наиболее тесно с первой главной компонентой коррелируют дескрипторы «ММВБ», «РТС», «дефолт» и «банкротство». Значения коэффициентов корреляции равны 0,95, 0,85, 0,72 и 0,79 соответственно.

Проверим, как построенный барометр соотносится с Индексом финансовых настроений (ИФН) Сбербанка РФ, который рассчитывается с мая 2009 г. Данный индикатор опирается на опросные данные, поэтому нужно посмотреть, насколько существенна корреляция схожих по смыслу, но заметно различающихся по методологии показателей финансовой конъюнктуры.

Коэффициент корреляции барометра финансовой конъюнктуры и ИФН равен  $-0,48$ . Теснота корреляционной связи двух индикаторов умеренная, но видно, что оба показателя в целом фиксируют одинаковый тренд изменения финансовой конъюнктуры (рис. 5). Увеличение ИФН (рост доверия к финансовой системе), как правило, совпадает с понижательной динамикой барометра.

Определенная «зашумленность» связи вызвана наличием в структуре барометра индивидуальных показателей, рост которых до ра-

<sup>20</sup> Избранная методика построения барометра финансовой конъюнктуры аналогична использованной в серии работ А. А. Френкеля, Н. Н. Райской и Я. В. Сергиенко, посвященных моделированию народно-хозяйственной конъюнктуры и инфляционных процессов в России (Райская Н. Н., Сергиенко Я. В., Френкель А. А. Использование интегральных индексов в анализе циклических изменений российской экономики // Вопросы статистики. 2009. № 12. С. 8—12). Выбор в пользу определения весов дескрипторов в интегральном показателе (барометре) через корреляционную матрицу обусловлен единообразным измерением данных — с помощью нормализованной частоты пользовательских запросов в Google.

**Сравнительная динамика барометра финансовой конъюнктуры в России и Индекса финансовых настроений Сбербанка РФ, май 2009 — июль 2011 г.**

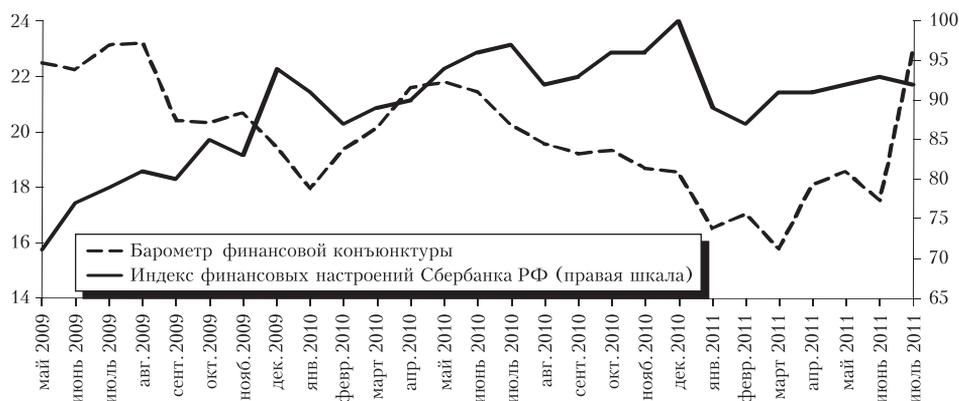


Рис. 5

зумных пределов свидетельствует об оживлении соответствующего сегмента финансовой системы («ПИФ», «акции»). Если анализировать взаимосвязь отдельных, близких по смыслу компонентов барометра и ИФН, то значения коэффициентов корреляции будут более существенными. В частности, между дескриптором «взять кредит» и компонентом ИФН «склонность к покупкам в кредит» он равен 0,6. Кроме того, присутствует довольно тесное корреляционное отношение между рассмотренным выше дескриптором «вклады (депозиты)» и компонентом ИФН «оценка целесообразности хранения сбережений в банках в ближайший год»: соответствующий коэффициент составил 0,72.

Закономерен вопрос о независимости ИФН и построенного барометра: не доминирует ли один из показателей другой. Для этого был реализован тест Гранджера для выявления возможной причинно-следственной связи этих индикаторов. С вероятностью 99% гипотеза, что один из них может выступить как предиктор динамики другого, была отвергнута. Таким образом, данные показатели автономны и не доминируют друг друга.

Следующий шаг — выяснение сравнительной эффективности двух показателей как индикаторов финансовой конъюнктуры. Здесь следует отметить, во-первых, что категория «финансовая конъюнктура» допускает различные трактовки. Под финансовой конъюнктурой можно понимать какой-либо один статистический показатель (например, динамику фондового индекса) или целую систему индикаторов, характеризующих различные сегменты финансовой системы и, возможно, государственных финансов. Индекс финансовых настроений Сбербанка и сконструированный барометр лишь отображают финансовую конъюнктуру. Адекватность таких отображений может изменяться в зависимости от того, что именно мы понимаем под этой конъюнктурой.

Во-вторых, сравнительный анализ возможен на небольшом временном интервале, поскольку ИФН рассчитывается лишь с мая 2009 г. Данное обстоятельство снижает надежность результатов применения весьма популярной в таких случаях техники векторных авторегрессионных моделей.

Не претендуя на комплексное преодоление упомянутых методологических сложностей, сравним, насколько барометр финансовой конъюнктуры и Индекс финансовых настроений Сбербанка улучшают качество эконометрических моделей абсолютного прироста депозитов физических лиц, задолженности по кредитам населению, объема выданных ипотечных жилищных кредитов, прироста совокупных активов российской банковской системы, темпов прироста индекса РТС. Этот набор показателей, на наш взгляд, в целом отражает состояние деловой активности в отечественной банковской системе и на финансовых рынках, хотя он не исчерпывающий. На примере этих финансовых переменных проанализируем, насколько ИФН и барометр финансовой конъюнктуры, построенный на основе пользовательских запросов в Google, повышают точность оценки динамики избранных показателей.

Для них изначально выбрана наиболее адекватная авторегрессионная модель (там, где возможно), в которую затем включались первые разности ИФН ( $D\_IFS$ ) и барометра финансовой конъюнктуры ( $D\_Bar$ ).

Отбор наиболее адекватных авторегрессионных моделей производился на базе двух показателей — скорректированного коэффициента детерминации и информационного критерия Акаике (AIC). При невозможности подбора авторегрессионных зависимостей использовались парные регрессии, при этом зависимые переменные и предикторы предварительно были приведены к стационарному виду. В большинстве случаев для этого хватило взятия первых разностей, то есть абсолютных приростов. Соответствующая проверка была проведена с помощью расширенного теста Дики—Фуллера и коинтеграционного теста Йохансена.

Ввиду ограниченности временного ряда Индекса финансовых настроений Сбербанка производились только внутривыборочные оценки вклада ИФН и нашего барометра финансовой конъюнктуры в прогнозирование перечисленных выше переменных. Преимущество этих переменных по сравнению с исходными моделями выражается в увеличении скорректированного коэффициента детерминации и уменьшении AIC. Соответствующие модели были построены на интервале май 2009 — июнь 2011 г. Результаты приведены в таблице 2.

Итак, в части банковских показателей перевес, впрочем, не абсолютный, на стороне Индекса финансовых настроений Сбербанка. Предложенный нами барометр помогает уточнить модели абсолютного прироста задолженности по кредитам физических лиц, а также динамики индекса РТС.

С учетом вариативности границ понятия «финансовая конъюнктура» подобные тесты могут быть продолжены. Однако имеющиеся результаты уже позволяют утверждать, что Индекс финансовых настроений и предложенный в настоящей статье барометр финансовой конъюнктуры, скорее, не конкурируют, а дополняют друг друга. Использование барометров финансовой конъюнктуры, построенных на статистике интернет-запросов, даст возможность уточнить результаты, полученные на основе опросных методов. В перспективе, при условии дальнейшего совершенствования процедуры отбора дескрипторов и нарастания «виртуализации» финансовых транзакций, такой подход может составить конкуренцию опросным методам оценки финансовой конъюнктуры.

**Вклад Индекса финансовых настроений Сбербанка РФ и барометра финансовой конъюнктуры в объяснение динамики ряда показателей развития банковской системы и фондового рынка России**

Показатель	Значимость* Индекса финансовых настроений Сбербанка	Значимость барометра финансовой конъюнктуры	Итоговая модель	Изменение скорректированного $R^2$ и AIC
Абсолютный** прирост депозитов физических лиц	+	-	$D\_g = 700538 + 0,24D\_g(-10) + 0,52D\_g(-12) + 14130D\_IFS$	Увеличение $R^2$ с 0,53 до 0,67; снижение AIC с 25,9 до 25,5
Абсолютный прирост задолженности по кредитам физических лиц	-	+	$D\_loan = 12378 + 0,86D\_loan(-1) + 12336D\_Bar$	Увеличение $R^2$ с 0,72 до 0,79; снижение AIC с 23,4 до 23,1
Абсолютный прирост объема выданных ипотечных жилищных кредитов	+	-	$Mort\_loan = 9698,6 + 0,8Mort\_loan(-4) + 1470,5D\_IFS$	Увеличение $R^2$ с 0,38 до 0,54; снижение AIC с 21,7 до 21,4
Абсолютный прирост активов банковской системы	+	-	$D\_bank\_assets = 276536 + 41627D\_IFS$	$R^2 = 0,17$
Темп прироста индекса РТС	-	+	$D\_RTS = 2 - 2,3D\_Bar$	$R^2 = 0,13$

\* Под значимостью ИФН и барометра финансовой конъюнктуры понимается значение  $p$ -value  $t$ -статистики коэффициента при этой переменной; «+» свидетельствует о статистической значимости переменной ( $p < 0,05$ ), «-» — об отсутствии значимости ( $p > 0,05$ ).

\*\* Использование темпов прироста зависимых и независимых переменных вместо абсолютных приростов не привело к повышению качества моделей.

Источник: рассчитано по данным Банка России и investfunds.ru.

\* \* \*

Проведенное исследование демонстрирует возможности использования статистики запросов пользователей Интернета в Google для оценки конъюнктуры российской финансовой системы. Показана перспективность использования этих данных для более эффективного моделирования динамики депозитов физических лиц. Был построен барометр финансовой конъюнктуры по широкому кругу дескрипторов, позволяющий лучше выявить масштаб *текущих* изменений в финансовой системе (nowcasting). Установлена статистически значимая линейная взаимосвязь барометра и его отдельных компонентов с Индексом финансовых настроений Сбербанка РФ, построенном по принципиально иной методологии. Разумеется, полученные результаты следует воспринимать как «пробный камень» в области финансово-конъюнктурных исследований российской экономики. Перспективным представляется использование статистики запросов в других поисковых системах (например, Яндекс), а также в социальных сетях и блогах (Живой журнал, Twitter, Facebook).