ВОПРОСЫ ТЕОРИИ

Гарантирующий и вероятностный подходы к выбору ставок штрафа и налогообложения

Закревский А.В., Токарев В.В.

Рассматривается проблема планирования государством долговременных управляющих воздействий – постоянных ставок налога и штрафа за сокрытие доходов. При этом распределение ресурсов налоговой службы между раскрытием и наказанием уклонений от уплаты налогов трактуется как оперативное управление, осуществляемое по алгоритмам построением для динамической модели в предыдущей статье [4]. В условиях неточного прогнозирования реакций налогоплательщиков и налоговой службы на величины ставок налога и штрафа используется гарантирующий и вероятностный подходы к планированию. В обоих случаях проблему планирования долговременных управляющих воздействий удалось свести к статическим задачам детерминированной минимизации, различающимся только способом исчисления критических возмущений. Сравниваются аналитические и численные результаты минимизации, обсуждаются их экономические особенности.

Показано, что с позиции налоговой службы оптимальна самая низкая ставка налога из априори оговоренного диапазона. Этот вывод справедлив независимо от использованного способа планирования, гарантирующего или вероятностно-гарантирующего, при самых общих естественных предположениях о поведенческих характеристиках.

Для оптимальной ставки штрафа за сокрытие доходов нет однозначного ответа: она может оказаться и на нижней, и на верхней границах разрешенного диапазона (редко внутри). Ответ зависит от соотношения скоростей уменьшения интенсивности потока сокрытий и падения эффективности работы налоговой службы при увеличении ставки штрафа.

Работа носит поисковый теоретический характер и не базируется пока на реальной количественной информации, а только на качественных наблюдаемых свойствах.

1. Введение

В настоящей работе продолжено изучение противодействия со стороны государства уклонениям от налогов, начатое в статье [4]. Там была построена динами-

Закревский А.В. – аспирант Московского физико-технического института.
Токарев В.В. – доктор физ.-мат. наук, профессор кафедры высшей математики ГУ ВШЭ.

Статья поступила в Редакцию в марте 2004 г.

ческая макромодель криминальной обстановки в экономике. Для этой модели была сформулирована задача оптимального распределения ресурсов налоговой службы между двумя видами ее деятельности: выявлением уклонений от налогов и их наказанием с целью минимизации суммарного объема невыявленных и ненаказанных сокрытий доходов налогоплательщиками.

Сформулированная задача оптимального управления на конечном отрезке времени была решена аналитически с помощью принципа максимума. Программа оптимального управления получилась релейной, типичной для задач, линейных по управлению: сначала все ресурсы используются для раскрытия уклонений от налогов, а затем, через какое-то время, они мгновенно переориентируются на исполнение наказаний. Оптимальный момент переключения найден как неявная функция от исходных данных задачи.

Релейная программа управления трудна для реализации, поэтому она была наилучшим образом аппроксимирована в классе программ с постоянным по времени распределением ресурсов. Были рассмотрены вопросы о вычислениях исходных данных, не все из которых непосредственно измеримы, по результатам доступных наблюдений за прошлым состоянием.

Во всех построениях, проделанных в [4], ставки налога и штрафа полагались фиксированными. Теперь они будут выбираться.

В странах со сложившейся рыночной экономикой налоговые ставки и штрафные санкции уже сформировались в результате многолетнего эмпирического подбора и на основе качественных теоретических рассуждений о немонотонной зависимости налоговых поступлений в бюджет от ставки налога (кривые Лафферовского вида).

В России же до сих пор идут горячие дискуссии о необходимости снижения налоговых ставок и о пользе или вреде жестких и нежестких наказаний. Проводятся также и многочисленные модельные исследования, краткий обзор которых был дан во введении к [4].

Словесная формулировка всей рассматриваемой проблемы такова. Государство хочет выбрать ставки налога и штрафа в пределах ограниченных диапазонов, чтобы при последующем оптимальном распределении ограниченного ресурса между раскрытием и наказанием уклонений от налогов был бы минимизирован суммарный объем невыявленных и ненаказанных сокрытий доходов к концу планового периода. Снижая по этому критерию криминогенность состояния экономики, государство одновременно обеспечивает высокий уровень возврата денежных средств в бюджет, пропорционально объему наказанных сокрытий.

Проблема анализируется в рамках динамической модели (13) из предшествующей статьи [4]. Модель построена на основе умозрительных качественно правдоподобных гипотез относительно поведенческих характеристик налогоплательщиков и налоговой службы. В модели учитывается, что затраты ресурсов требуются не только на выявление скрытых доходов, но и на реализацию штрафных санкций по выявленным нарушениям налогового законодательства.

Кроме того, и налогоплательщики, и налоговая служба реагируют в модели на предысторию процесса, которая характеризуется текущими объемами выявленных и невыявленных, но пока не наказанных сокрытий. С ростом этих объемов увеличивается интенсивность потока новых сокрытий, но одновременно повышаются скорости раскрываемости и реализации наказаний. Какие из этих тенденций окажутся преобладающими, зависит от ставок налога и штрафа.

Сначала обе ставки фиксировались, и тогда с помощью принципа максимума в [4] была решена динамическая задача оптимального распределения ограниченного ресурса между розыском и наказаниями. В результате минимизируемый критерий удалось представить в виде аналитической функции от трех постоянных ключевых параметров динамической модели. Это максимальный темп роста новых сокрытий (характеристика налогоплательщиков), максимальный темп выявления сокрытий (характеристика розыскных возможностей налоговой службы) и максимальный темп реализации наказаний (характеристика возможностей исполнения наказаний).

Качественные посылки относительно зависимости ключевых параметров от ставок налога и штрафа, как правило, не вызывают возражений. Чем больше налоговая ставка, тем больше скрывается доходов, тем тщательнее конспирируются сокрытия, тем более активна коррупция, и поэтому эффективность розыска сокрытий и реализации наказаний падает. Ужесточение штрафных санкций сдерживает сокрытия, но их снова труднее раскрывать и наказывать.

Однако формульная реализация одних и тех же качественных гипотез неоднозначна, а от нее могут сильно зависеть итоговые выводы. В настоящей статье делается естественный шаг на пути преодоления этой трудности в условиях недостаточной информированности. Вид поведенческих формул задается, но в них предусматривается достаточное число параметров, которые точно не известны, а для выбора ставок налогов и штрафов используются гарантирующий или вероятностный подходы. Кроме того, часть выводов оказалась независимой от формульной реализации поведенческих характеристик.

Конечно, такие построения еще не закрывают проблему адекватности модельного описания. Требуется дальнейшая работа по сбору статистических данных и идентификации поведенческих характеристик. Второе направление, дополняющее упомянутое выше, – это организация схемы управления, подстраивающегося под наблюдаемую реальную ситуацию.

2. Общие положения гарантирующего планирования долговременных управлений

Формализация проблемы гарантирующего планирования в условиях неопределенности, развитая в работах [1, 6, 7, 8], здесь адаптируется для последующего использования в рассматриваемой задаче противодействия сокрытию доходов на этапе планирования ставок штрафа и налогообложения.

Основные понятия:

ξ – вектор *возмущений*, т.е. неконтролируемых и неточно прогнозируемых параметров и функций времени (начальный объем x_0 скрытых доходов, λ – темп их несдерживаемого роста, $\omega_{1,2}$ – максимальные темпы выявления сокрытий и реализации наказаний);

$\Xi = \{\xi\}$ – прогнозируемое априори, к моменту планирования, *множество ожидаемых возмущений* (задание вероятности распределений на Ξ пока не предполагается, считается только, что возмущения никогда не выходят за пределы Ξ и любое возмущение $\xi \in \Xi$ может реализоваться);

$w \in W$ – *долговременные управления*, планируемые заранее по априорной информации Ξ о возмущениях ξ (w_1 – ставка штрафа, w_2 – ставка налогообложения, неизменные во времени);

u – *оперативные управления*, выбираемые в процессе функционирования по некоторому правилу $\tilde{U}$ на основе текущей информации о возмущениях и о состоянии управляемого объекта при известных долговременных управлениях w (u – доля сотрудников в налоговой службе, используемых для выявления скрытых доходов, $(1-u)$ – доля сотрудников, занимающихся реализацией наказаний);

$U(w,\xi)$ – *множество допустимых оперативных управлений*, подверженное в общем случае действию возмущений и зависящее от долговременных управлений (задается всеми ресурсными и целевыми ограничениями в силу уравнений динамики объекта – см. (13) из [4]);

$J(w,u,\xi)$ – максимизируемый скалярный *критерий качества* управления, тоже зависящий от возмущений, быть может и неявно (в задаче (13) из [4] минимизируется суммарный объем невыявленных и ненаказанных сокрытий в конечный момент времени).

В применяемом далее общем принципе гарантированного результата [1] явно отражаются, как это сделано в [6–8], две существенные особенности, присущие большинству задач экономического управления в условиях неопределенности:

- от возмущений зависит не только критерий качества, но и множество допустимых оперативных управлений, как в играх с запрещенными ситуациями, лишь только бегло рассматриваемых в теории игр [2];
- при планировании долговременных управлений возмущения не известны, прогнозируется только множество ожидаемых возмущений, но при этом полезно принимать во внимание правило оперативного реагирования на будущие реализации возмущений.

Принцип гарантированного результата – это модель осторожного, безрискового, поведения в условиях неопределенности. Решения должны приниматься так, чтобы при любой реализации возмущений из прогнозируемого множества $\xi \in \Xi$ была бы обеспечена допустимость результирующего управления, а его качество J оказалось бы не хуже априорной оценки J^0, которая была бы при этом максимально возможной.

Не все типы управленческих действий укладываются в такую схему. Возможны и рискованные решения, для которых допускается некоторая приемлемая вероятность нереализуемости ожиданий (см. разд. 4). Но тогда нужен прогноз вероятности распределения возмущений ξ по множеству Ξ. Если же этот прогноз отсутствует, то принцип гарантированного результата является, пожалуй, единственным осмысленным основанием для принятия управленческих решений, за исключением неформализуемых безрассудных действий.

Ниже приводится *формульная расшифровка* принципа гарантированного результата на этапе планирования долговременных управлений w при известном правиле $\tilde{U}(w,\xi)$ оперативного реагирования u на будущие возмущения ξ.

Во-первых, план w не может зависеть от ξ и вместе с тем должен обеспечить допустимость оперативного управления $u=\tilde{U}(w,\xi)$ во всех возможных ситуациях. Такие планы называются *гарантированно допустимыми* и их множество W^0 определяется следующим образом:

$$W^0=\{w\in W:\forall\xi\in\Xi\quad u=\tilde{U}(w,\xi)\in U(\xi)\}. \tag{1}$$

Если гарантированно допустимые планы существуют и неединственны, то из них можно выбрать *оптимальный гарантирующий план* $w^0\in W^0$, обеспечивающий максимум гарантированной оценки качества

$$\inf_{\xi\in\Xi} f(w^0,\xi)=\max_{w\in W^0}\left[\inf_{\xi\in\Xi} f(w,\xi)\right]=f^0,\text{ где } f(w,\xi)=J(w,\tilde{U}(w,\xi),\xi). \tag{2}$$

Выйти за пределы множества (1) нельзя, ибо тогда в силу его определения найдется ожидаемое возмущение $\xi'\in\Xi$, которое сделает управление недопустимым $u=\tilde{U}(w,\xi')\,\bar{\in}\,U(\xi')$.

Отклонение же от условия (2) нерационально, так как любой гарантированно допустимый план $w'\in W^0$, отличный от необязательно единственного плана w^0 из (2), не принесет столь высокой гарантированной оценки качества.

Можно, конечно, более оптимистично оценивать качество управления, чем по принципу гарантированного результата (1), (2). Например, можно усреднить критерий $J(w,u,\xi)$ по множеству Ξ, считая возмущения ξ равновероятными. Но тогда при некоторых возмущениях из Ξ реализация критерия J может оказаться хуже ожидаемой оценки, что вряд ли приемлемо, по крайней мере, для уникальных операций.

3. Оптимальные гарантирующие ставки штрафа и налогообложения

Основной источник возмущений в проблеме противодействия уклонениям от налогов − это плохое знание на количественном уровне поведенческих характеристик налогоплательщиков и эффективности работы налоговой службы.

В принятом здесь описании такие неопределенности сказываются на трех безразмерных коэффициентах (9а) из [4], от которых зависит решение (33), (36) или (42), (43) из [4] задачи оперативного управления:

λ − максимальный темп роста невыявленных сокрытий (при отсутствии сдерживающих управлений),

ω_1 − максимальный темп выявления сокрытий (при использовании всех ресурсов только для поиска),

ω_2 − максимальный темп реализации наказаний (при использовании всех ресурсов только для наказания).

На этапе оперативного управления не требовалось знания зависимости этих коэффициентов от долговременных управляющих параметров $w_{1,2}$ − ставок налога и штрафа. Там было достаточно возможности их вычисления по результатам измерений наблюдаемых величин, как показано в разд. 7 из [4]. Однако на этапе планирования долговременных управлений без знания, хотя бы и неточного, таких зависимостей не обойтись.

Здесь принимается простейший вариант, соответствующий качественным гипотезам 1–11 и формулам (3)–(5) из [4]:

$$(3)\qquad \lambda=\alpha_1 w_1\left(1-\alpha_2 w_2\right)T,\quad \omega_1=\beta_1 T/\left(w_1+w_2\right),\quad \omega_2=\beta_2 T/\left(w_1+w_2\right),\ \text{где}\ \ \alpha_2 w_2\le 1,$$

т.е. предполагается, что темп роста сокрытий λ линейно увеличивается с ростом ставки налога w_1 и линейно падает с ростом ставки штрафа w_2, а темпы выявления сокрытий ω_1 и реализации наказаний ω_2 обратно пропорциональны сумме ставок w_1+w_2.

Однако коэффициенты пропорциональности в соотношениях (3) на этапе выбора ставок $w_{1,2}$ известны неточно. Прогнозируются только границы $\underline{a}_{1,2},\overline{a}_{1,2}$ и $\underline{b}_{1,2},\overline{b}_{1,2}$ диапазонов их возможных значений:

$$(4)\qquad \underline{a}_i\le\alpha_i\le\overline{a}_i,\quad \underline{b}_i\le\beta_i\le\overline{b}_i,\quad i=1,2.$$

Помимо поведенческих неопределенностей (4) при выборе ставок налогообложения и штрафа нужно учитывать, что единожды назначенные ставки будут использоваться в различных ситуациях: при разных относительных объемах невыявленных сокрытий x_0 из прогнозируемого диапазона

$$(5)\qquad \underline{x}_0\le x_0\le\overline{x}_0.$$

Природа неопределенностей (4) и (5) различна, но для реализуемой далее общей процедуры (1), (2) эти различия несущественны – все параметры из (4), (5) играют роль возмущений $\xi=\left(\alpha_1,\alpha_2,\beta_1,\beta_2,x_0\right)$. Информации о связях между компонентами вектора возмущений здесь не предполагается. По этой причине приходится считать, что возмущения могут принимать независящие друг от друга любые значения из диапазонов (4), (5).

Множество допустимых оперативных управлений $\left(u,v\right)$, задающих интенсивность розыска u и реализации наказаний v, в задаче (13) из [4] от возмущений не зависит, поскольку ограничения на текущие значения этих управлений $u,v\ge 0,\ u+v\le 1$ возмущений не содержат, а конечные условия на фазовые координаты в задаче отсутствуют. Ее решение существует при любых величинах x_0,λ,ω_1 и ω_2, так что надобность в построении множества (1) гарантированно допустимых ставок $w_{1,2}$ отпадает.

Точная гарантированная оценка критерия J находится аналитически, только вместо точной нижней грани по возмущениям, записанной в формуле (2), нужно отыскивать верхнюю грань, так как по управлениям критерий минимизируется.

И для релейного, и для постоянного оперативного управления зависимость критерия J от четырех безразмерных параметров $\eta=\left(x_0,\lambda,\omega_1,\omega_2\right)$, задаваемая соотношениями (33), (36) или (42), (43) из [4], непрерывна. В свою очередь все эти параметры в силу (9а) из [4] и (3) непрерывно зависят от возмущений $\xi = \left(\alpha_{1,2}, \beta_{1,2}, x_0\right)$.

Критерий J содержит еще один параметр $s = t_*$ для (33), (36) из [4] или $s = u^0$ для (42), (43) из [4]. Параметр s определяется как неявная функция $\Phi\left(s,\eta\right) = 0$ от параметров η.

В результате итоговая зависимость критерия от возмущений ξ

(6) $f\left(w,\xi\right) = J\left(\eta\left(w,\xi\right), s\right)$ при $\Phi\left(s,\eta\right) = 0$, где $\eta = \left(x_0, \lambda, w_1, w_2\right)$, где $s = t_*$ или $s = u^0$

также оказывается непрерывной. Так что на ограниченном замкнутом и непустом множестве (4), (5) точная верхняя грань по ξ функции (6) достигается.

По всем возмущениям (4), (5) функция (6) монотонна. Этот факт устанавливается из анализа знаков ее частных производных

$$\frac{\partial f}{\partial \xi_i} = \sum_{j=1}^{4}\left(\frac{\partial J}{\partial \eta_j}\frac{\partial \eta_j}{\partial \xi_i}\right) + \frac{\partial J}{\partial s}\sum_{j=1}^{4}\left(\frac{\partial s}{\partial \eta_j}\frac{\partial \eta_j}{\partial \xi_i}\right),\ i = 1,...,5. \tag{7}$$

Производная $\partial J/\partial s \equiv 0$, так как в одной части пространства параметров η согласно (33), (36) или (42), (43) из [4] $s \equiv 0$, а в оставшейся, где $s > 0$, величина параметра s выбирается из условия минимума критерия J, откуда и получается $\partial J/\partial s = 0 \Leftrightarrow \Phi\left(s,\eta\right) = 0$.

Параметр s назначается на этапе оперативного управления, когда истинные значения неизвестных заранее параметров η восстановлены по результатам наблюдений, как об этом говорилось в разделе 7 из [4].

Таким образом, в формуле (7) исчезает последняя сумма, и остается только

$$\frac{\partial f}{\partial \xi_i} = \sum_{j=1}^{4}\left(\frac{\partial J}{\partial \eta_j}\frac{\partial \eta_j}{\partial \xi_i}\right),\ i = 1,...,5. \tag{8}$$

Не воспроизводя здесь весьма громоздких формальных выкладок, ограничимся констатацией ожидаемых промежуточных результатов:

$$\frac{\partial J}{\partial \eta_1} = \frac{\partial J}{\partial x_0} > 0,\ \frac{\partial J}{\partial \eta_2} = \frac{\partial J}{\partial \lambda} > 0,\ \frac{\partial J}{\partial \eta_3} = \frac{\partial J}{\partial \omega_1} \le 0,\ \frac{\partial J}{\partial \eta_4} = \frac{\partial J}{\partial \omega_2} < 0. \tag{9}$$

Они означают, что чем больше начальный объем x_0 невыявленных сокрытий и темп их роста λ, тем хуже. Но чем выше эффективность розыска ω_1 и реализации наказаний ω_2, тем лучше.

Соотношения (8), (9) позволяют установить знакоопределенность производных по параметрам $\alpha_{1,2}$ и $\beta_{1,2}$ в поведенческих характеристиках (3) и отыскать наихудшие, т.е. максимизирующие значения этих параметров из диапазонов (4) их возможных значений:

$$\begin{aligned}
&\frac{\partial f}{\partial \alpha_1}=\frac{\partial J}{\partial \lambda}w_1\left(1-\alpha_2 w_2\right)T\geq 0 \text{ при } \alpha_2 w_2\leq 1\Rightarrow \max_{\alpha_1\in\left[\underline{a}_1,\overline{a}_1\right]} f\left(\alpha_1,\ldots\right)=f\left(\overline{a}_1,\ldots\right),\\
&\frac{\partial f}{\partial \alpha_2}=-\frac{\partial J}{\partial \lambda}\alpha_1 w_1 w_2 T<0\Rightarrow \max_{\alpha_2\in\left[\underline{a}_2,\overline{a}_2\right]} f\left(\alpha_2,\ldots\right)=f\left(\underline{a}_2,\ldots\right),\\
&\frac{\partial f}{\partial \beta_1}=\frac{\partial J}{\partial \omega_1}\frac{T}{w_1+w_2}\leq 0\ \Rightarrow \max_{\beta_1\in\left[\underline{b}_1,\overline{b}_1\right]} f\left(\beta_1,\ldots\right)=f\left(\underline{b}_1,\ldots\right),\\
&\frac{\partial f}{\partial \beta_2}=\frac{\partial J}{\partial \omega_2}\frac{T}{w_1+w_2}<0\ \Rightarrow \max_{\beta_2\in\left[\underline{b}_2,\overline{b}_2\right]} f\left(\beta_2,\ldots\right)=f\left(\underline{b}_2,\ldots\right).
\end{aligned} \tag{10}$$

Естественно, что наихудшими получаются самая острая реакция производителей на ставку налогообложения $\alpha_1=\overline{a}_1$ и вялое реагирование на ставку штрафа $\alpha_2=\underline{a}_2$, а также предельное усиление скрытности, снижающее эффективность работы налоговой службы $\beta_{1,2}=\underline{b}_{1,2}$.

Монотонно воздействует на критерий и начальный объем x_0 невыявленных сокрытий

$$\frac{\partial f}{\partial x_0}=\frac{\partial J}{\partial x_0}>0\Rightarrow \max_{x_0\in\left[\underline{x}_0,\overline{x}_0\right]} f\left(x_0,\ldots\right)=f\left(\overline{x}_0,\ldots\right), \tag{11}$$

так что критичной для этой компоненты вектора возмущений оказывается верхняя граница диапазона (5): $x_0=\overline{x}^0$.

Найденные в (10), (11) наихудшие значения возмущений ξ^* не зависят от плана w:

$$\exists\xi^*=\left(\overline{a}_1,\underline{a}_2,\underline{b}_1,\underline{b}_2,\overline{x}_0\right)\in\Xi:\forall w\in W\quad f\left(w,\xi^*\right)=\max_{\xi\in\Xi} f\left(w,\xi\right). \tag{12}$$

Если при таких значениях возмущения достижим еще и минимум f по w

$$\exists w^*\in W: f\left(w^*,\xi^*\right)=\min_{w\in W} f\left(w,\xi^*\right), \tag{13}$$

то свойств (12), (13) достаточно для наличия седловой точки. Таковой является пара $\left(w^*,\xi^*\right)$, поскольку для нее в силу (12), (13) оказывается выполненным определение седловой точки

$$\exists\left(w^*,\xi^*\right)\in W\cdot\Xi:\min_{w\in W} f\left(w,\xi^*\right)=f^*=f\left(w^*,\xi^*\right)=\max_{\xi\in\Xi} f\left(w^*,\xi\right). \tag{14}$$

С позиций теории управления наличие седловой точки (14) означает, что используемый здесь простейший программный способ управления $w\left(\xi\right)\equiv\text{const}$ дает такую же гарантированную оценку качества f^*, как идеальное управление по точной информации о возмущениях. Это заключение следует из доказываемого в теории антагонистических игр [1] необходимого и достаточного условия седловой точки:

(15) $$\exists f^*:\ \min_{w\in W}\left[\sup_{\xi\in\Xi} f\left(w,\xi\right)\right] = f^* = \max_{\xi\in\Xi}\left[\inf_{w\in W} f\left(w,\xi\right)\right].$$

Таким образом, в рассматриваемой задаче все разумные способы управления оказываются неразличимы по гарантированной оценке качества. В этих условиях предпочтение следует отдать программному способу управления как самому простому. Однако высказанную рекомендацию нужно смягчить, так как интерес представляют значения критерия не только в критических точках ξ^*. А в некритических ситуациях $\xi \neq \xi^*$ проигрыш в критерии программного управления по сравнению с идеальным

$$f\left(w^*,\xi\right) - \inf_{w\in W} f\left(w,\xi\right)$$

может оказаться существенным даже при наличии равенства (15).

Перейдем теперь непосредственно к выбору *оптимального гарантирующего плана* w^*. В силу свойства (12) независимости критических возмущений ξ^* от плана w и в гарантирующей постановке, и в детерминированной, где возмущение ξ как-то фиксируется заранее (не обязательно $\xi = \xi^*$), оптимальный план w^0 находится в результате решения обычной задачи минимизации с фиксированным параметром ξ:

(16) $$w^0\left(\xi\right): f\left(w^0\left(\xi\right),\xi\right) = \min_{w\in W} f\left(w,\xi\right) \text{ при } \xi = \mathrm{fix} \in \Xi.$$

Также строится и упоминавшееся ранее идеальное управление, но оно предполагает точное знание возмущения ξ и возможность смены управления w при изменениях ξ.

Оптимальный гарантирующий план берется из семейства решений задачи (16) как $w^* = w^0\left(\xi^*\right)$ и сохраняется постоянным для любых возмущений $\xi \in \Xi$.

Вектор планов $w = \left(w_1, w_2\right)$ состоит из двух компонент: w_1 − ставка налога и w_2 − ставка штрафа. Допустимое множество W задается как прямое произведение двух отрезков

(17) $$W = \left[\underline{w}_1, \overline{w}_1\right]\cdot\left[\underline{w}_2, \overline{w}_2\right].$$

Что касается *ставки налога* w_1, то с позиции налоговой службы ее всегда выгодно назначать минимально допустимой:

(18) $$w_1^0 \equiv \underline{w}_1 \quad \forall\xi \in \Xi,\ \ \forall w_2.$$

Этот вывод справедлив при довольно общих предположениях о поведенческих характеристиках налогоплательщиков и налоговой службы

(19) $$\frac{\partial\lambda}{\partial w_1} > 0,\ \ \frac{\partial\omega_i}{\partial w_1} < 0,\ i = 1,2,$$

которым удовлетворяют и применяемые здесь функции (3).

В самом деле,

$$\frac{\partial f}{\partial w_1} = \frac{\partial J}{\partial \lambda}\frac{\partial \lambda}{\partial w_1} + \sum_{i=1}^{2}\frac{\partial J}{\partial \omega_i}\frac{\partial \omega_i}{\partial w_1} > 0,$$

так как согласно (9) $\partial J/\partial\lambda > 0$, а $\partial J/\partial\omega_i \le 0$, что и соответствует с учетом (19) положительности полной производной.

Полученный результат (18) вполне естественный. Ведь уменьшение налоговой ставки w_1 снижает интенсивность потока сокрытий λ и увеличивает эффективность работы налоговой службы, т.е. увеличивает интенсивность розыска ω_1 и наказаний ω_2.

Однако проблема рационального налогообложения такими соображениями отнюдь не исчерпывается. Здесь нужно принимать в рассмотрение и объем поступлений в государственный бюджет, и воздействие на инвестиционный климат, и социальные аспекты. Из такого комплексного рассмотрения и должны воспоследовать границы $\underline{w}_1, \overline{w}_1$ свобод налоговой ставки, доступных для выбора налоговой службой. Ее предпочтения, как здесь установлено, – нижняя граница $\underline{w}_1$ допустимого диапазона.

Выбор ставки штрафа w_2 представляется всецело прерогативой налоговой службы, и здесь оптимальное решение заранее не очевидно.

С ужесточением штрафов интенсивность потока сокрытий λ уменьшается, но одновременно с этим падают эффективность розыска ω_1 и наказаний ω_2. Так что знак полной производной

$$\frac{\partial f}{\partial w_2} = \frac{\partial J}{\partial \lambda}\frac{\partial \lambda}{\partial w_2} + \sum_{i=1}^{2}\frac{\partial J}{\partial \omega_i}\frac{\partial \omega_i}{\partial w_2}$$

зависит от конкретизации поведенческих характеристик налогоплательщиков $\lambda(w_2)$ и налоговой службы $\omega_{1,2}(w_2)$ и может меняться по w_2.

Численные расчеты, проведенные для семейства поведенческих характеристик (3), иллюстрируемых рис. 1а, показали, что в широком диапазоне назначения параметров семейства итоговая зависимость $f(w_2)$ получается немонотонной и выпуклой вверх. Типичный ее пример представлен на рис. 1б.

Минимумы таких функций, непрерывность которых была установлена прежде, достигаются либо на левой $\underline{w}_2$, либо на правой $\overline{w}_2$ границах допустимых значений их аргумента

$$\text{(20a)} \qquad \min_{w_2 \in [\underline{w}_2, \overline{w}_2]} f(w_2) = \min\{f(\underline{w}_2); f(\overline{w}_2)\}.$$

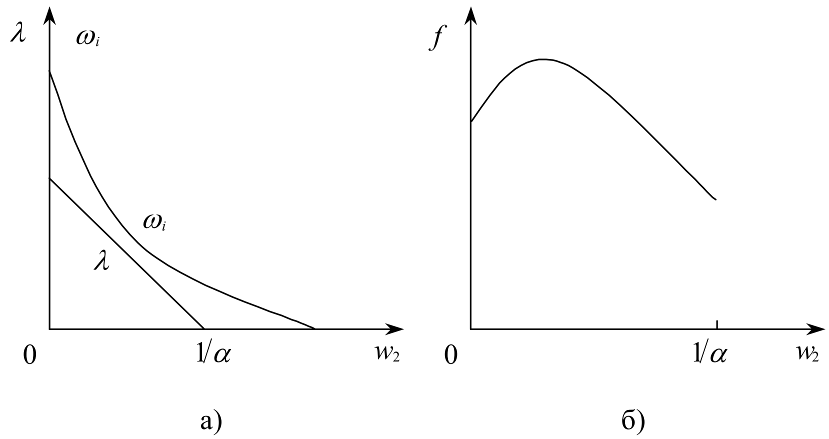

Рис. 1. Пример поведенческих характеристик (3)
и соответствующая ему зависимость $f(w_2)$ минимизируемого критерия от ставки штрафа w_2

Когда в (20а) случается равенство $f(\underline{w}_2)=f(\overline{w}_2)$, то предпочтение следует отдать меньшей ставке штрафа $\underline{w}_2$, с тем чтобы ослабить побудительные причины для коррупции налоговой службы.

Если диапазон изменения аргумента w_2 не стеснен ничем, кроме условия неотрицательности λ из (3), то минимум f всегда достигается на правой границе

$$\min_{w_2\in[0,1/\alpha]} f(w_2)=f(1/\alpha). \tag{20б}$$

Это и понятно, ведь согласно (3) $\lambda=0$, а $\omega_{1,2}>0$ при $w_2=1/\alpha$.

Однако в столь широком диапазоне искать минимум нельзя, поскольку в (3) принята линейная аппроксимация реальной зависимости $\lambda(w_2)$, плохо отражающая действительность вблизи $w_2=1/\alpha$. Так что в рамках (3) спор между сторонниками мягких и жестких наказаний нельзя считать разрешенным.

Тем не менее некоторые общие заключения можно сделать, не прибегая к конкретизации поведенческих характеристик. Для наглядности сократим число характеристик до двух, положив равными эффективности розыска и реализации наказаний $\omega_1=\omega_2=\omega$, так чтобы стала возможна геометрическая иллюстрация на плоскости λ − интенсивность сокрытий, ω − эффективность налоговой службы.

На плоскости (λ,ω) изобразим линии уровня минимизируемого критерия $J(\lambda,\omega)=\text{const}$ (рис. 2). Проведенные расчеты показали, что линии уровня представляют собой семейство почти параллельных прямых. Их наклон положительный, так как для компенсации большей интенсивности сокрытий λ, чтобы сохранить величину критерия, требуется увеличение эффективности ω работы налоговой службы.

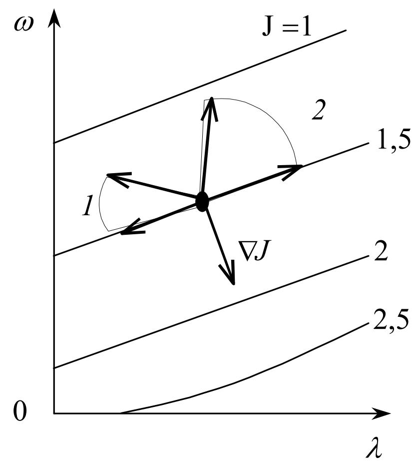

Рис. 2. Линии уровня минимизируемого критерия $J = \mathrm{const}$
и конусы *1*, *2* локально улучшающих смещений (заштрихованы)

Имея в распоряжении картину линий уровня $J = \mathrm{const}$, можно классифицировать поведенческие характеристики по их влиянию на оптимальную ставку штрафа, не прибегая к аналитическому заданию, подобному (3).

Локальная классификация основывается на вычислении знака производной критерия J по различным направлениям смещения из некоторой фиксированной точки $\lambda^0 = \lambda\left(w_2^0\right)$, $\omega^0 = \omega\left(w_2^0\right)$ в близкую к ней при малом изменении Δw_2 ставки штрафа

$$\lambda = \lambda\left(w_2^0 + \Delta w_2\right) \approx \lambda^0 + \left(\frac{\partial \lambda}{\partial w_2}\right)^0 \Delta w_2, \quad \omega^1 = \omega\left(w_2^0 + \Delta w_2\right) \approx \omega^0 + \left(\frac{\partial \omega}{\partial w_2}\right)^0 \Delta w_2 \text{ при} \left|\Delta w_2\right| << 1.$$

При таком векторе смещения $\left(\Delta\lambda, \Delta\omega\right)^{\mathrm{T}}$ критерий получит следующее приращение:

$$\Delta J \approx \left(\nabla J\right)^0 \frac{\Delta\lambda}{\Delta\omega} = \left[\left(\frac{\partial J}{\partial \lambda}\right)^0 \left(\frac{\partial \lambda}{\partial w_2}\right)^0 + \left(\frac{\partial J}{\partial \omega}\right)^0 \left(\frac{\partial \omega}{\partial w_2}\right)^0\right] \Delta w =$$

$$= \left[e_\lambda \left(\frac{\partial \lambda}{\partial w_2}\right)^0 + e_\omega \left(\frac{\partial \omega}{\partial w_2}\right)^0\right] h\left(\lambda^0, \omega^0\right) \Delta w_2, \text{ где } \nabla J = \left(e_\lambda, e_\omega\right) h, \ \left|e\right| = 1.$$

Длина h вектора ∇J переменная, а его направление почти неизменно, по расчетам оно близко к $e_\lambda = 0{,}86$, $e_w = -0{,}5$.

Тогда для уменьшения критерия $\Delta J < 0$ нужно назначить такие приращения Δw_2 ставки штрафа:

$$\text{(21)}\quad \begin{aligned} &\text{1)}\ \ \Delta w_2 > 0,\ \text{если}\ \left(\frac{\partial\lambda}{\partial w_2}\right)^0 < -\frac{e_\omega}{e_\lambda}\left(\frac{\partial\omega}{\partial w_2}\right)^0 \approx 1,72\left(\frac{\partial\omega}{\partial w_2}\right)^0, \\ &\text{1)}\ \ \Delta w_2 < 0,\ \text{если}\ \left(\frac{\partial\lambda}{\partial w_2}\right)^0 > -\frac{e_\omega}{e_\lambda}\left(\frac{\partial\omega}{\partial w_2}\right)^0 \approx 1,72\left(\frac{\partial\omega}{\partial w_2}\right)^0, \end{aligned}$$

где $\partial\lambda/\partial w_2 < 0,\ \partial\omega/\partial w_2 < 0$.

Если поведенческие характеристики $\lambda(w_2)$ и $\omega(w_2)$ таковы, что ни одно из неравенств в (21) не выполняется, то ставка штрафа w_2^0 локально неулучшаема.

Конусы локально улучшающих смещений (21) выделены на рис. 2 штриховкой. Без штриховки остались конусы невозможных и неулучшающих смещений. Геометрический смысл условий (21) вполне прозрачен: только те из возможных смещений могут локально улучшить анализируемую точку (λ^0,ω^0), которые ведут к линиям уровня с меньшими значениями критерия, чем $J(\lambda^0,\omega^0)$.

Для *глобальной классификации* поведенческих характеристик используется геометрический способ решения задач на условный экстремум. На плоскости (λ,ω) строится параметрически задаваемая поведенческая кривая $\omega = \varphi(\lambda)$, где $\lambda = \lambda(w_2)$, $\omega = \omega(w_2)$, для установленного диапазона изменения параметра $w_2 \in [\underline{w}_2, \overline{w}_2]$ (в пределе от 0 до $+\infty$). На этой кривой ищется точка касания или пересечения с линией уровня критерия $J(\lambda,\omega) = c = \mathrm{const}$, индексированной минимальной константой c.

Заранее заготовить ответ в конструктивном виде для всех мыслимых видов поведенческих кривых $\omega = \varphi(\lambda)$ невозможно, несмотря на то, что по естественным гипотезам для них выполняется свойство монотонного возрастания $\varphi'(\lambda) > 0$. Однако такие заготовки удается сделать для трех распространенных классов функций $\varphi(\lambda)$, что иллюстрируется на рис. 3а–3в. Там линии уровня $J = c$ показаны тонкими линиями, а характерные примеры функций $\omega = \varphi(\lambda)$ – жирными. Стрелками указаны направления роста параметра w_2. Звездочками отмечены положения минимума J.

Для *линейных функций* $\varphi(\lambda)$ (рис. 3а):

$$\text{(22a)}\quad \min_{w_2\in[\underline{w}_2,\overline{w}_2]} J = \begin{cases} J(\overline{w}_2),\ \text{если}\ \varphi'(\lambda) < -e_\lambda/e_\omega \approx 1,72\ \ (\textit{прямая 1}), \\ J(\underline{w}_2) = J(\overline{w}_2),\ \text{если}\ \varphi'(\lambda) = -e_\lambda/e_\omega \approx 1,72\ \ (\textit{прямая 2}), \\ J(\underline{w}_2),\ \text{если}\ \varphi'(\lambda) > -e_\lambda/e_\omega \approx 1,72\ \ (\textit{прямая 3}). \end{cases}$$

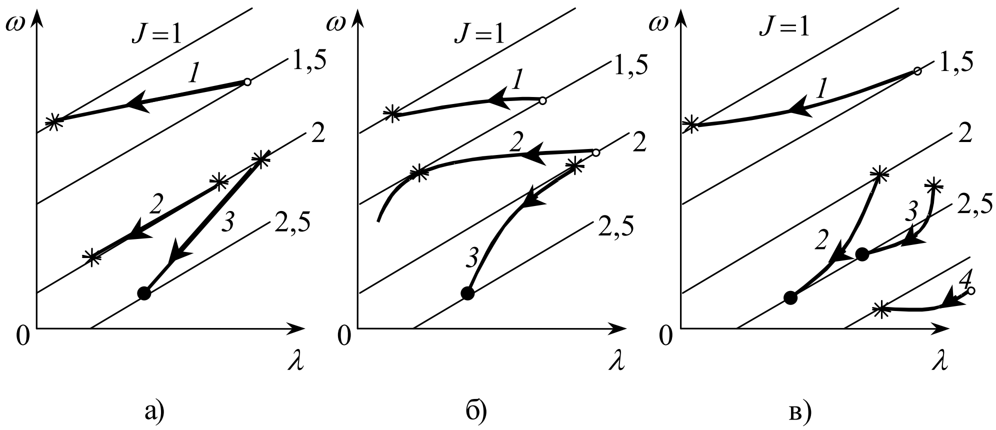

Рис. 3. Возможные положения минимума J для поведенческих кривых $\omega = \varphi(\lambda)$ с разными направлениями выпуклости

Для *выпуклых вверх функций* $\varphi(\lambda)$ (рис. 3б):

$$\min_{w_2 \in [\underline{w}_2, \overline{w}_2]} J = \begin{cases} J(\overline{w}_2), \text{ если } \varphi'(\lambda(\overline{w}_2)) < 1{,}72 \ (\textit{кривая 1}), \\ J(w_2^*), \text{ если } \varphi'(\lambda(\overline{w}_2)) \geq \varphi'(\lambda(w_2^*)) = 1{,}72 \geq \varphi'(\lambda(\underline{w}_2)) \ (\textit{кривая 2}), \\ J(\underline{w}_2), \text{ если } \varphi'(\lambda(\overline{w}_2)) > 1{,}72 \ (\textit{кривая 3}). \end{cases} \tag{22б}$$

Для *выпуклых вниз функций* $\varphi(\lambda)$ (рис. 3в):

$$\begin{aligned} &\min_{w_2 \in [\underline{w}_2, \overline{w}_2]} J = \min\{J(\underline{w}_2); J(\overline{w}_2)\} = J^*; \\ &\text{так, если } \varphi'(\lambda(\underline{w}_2)) \leq 1{,}72, \text{ то } J^* = J(\overline{w}_2) \ (\textit{кривая 1}), \\ &\text{если } \varphi'(\lambda(\overline{w}_2)) \geq 1{,}72, \text{ то } J^* = J(\underline{w}_2) \ (\textit{кривая 2}), \\ &\text{если } \varphi'(\lambda(\underline{w}_2)) \geq 1{,}72 \geq \varphi'(\lambda(\overline{w}_2)), \text{ то } J^* = J(\underline{w}_2) \ (\textit{кривая 3}) \\ &\text{или } J(\overline{w}_2) \ (\textit{кривая 4}). \end{aligned} \tag{22в}$$

Проведенный геометрический анализ подтвердил возможность обоих граничных минимумов, обнаруженную ранее в численных расчетах. Добавилась еще возможность внутреннего минимума (кривая *2* на рис. 3б).

4. Общие положения вероятностного планирования

Желанием улучшить априорную оценку качества управления по сравнению с гарантированной оправдываются вероятностные подходы к выбору управления. Однако их реализация требует дополнительных усилий. Нужно собрать и обработать статистику возмущений для исчисления вероятности различных реализаций. К тому же использование вероятностных управлений сопряжено с риском получения значений критерия качества, худших, чем его прогнозируемая оценка, и даже, вообще говоря, худших, чем гарантированная оценка.

Окончательное решение должно приниматься оперирующей стороной с учетом отмеченных особенностей вероятностного планирования.

Априорная информация о возмущениях:

$\Xi = \{\xi^1, ..., \xi^m\}$ – множество ожидаемых возмущений ξ^j (для простоты полагается пока конечным);

$\mu_j = P\{\xi = \xi^j\} \in [0;1]$, $j = 1,...,m$ – вероятность реализации каждого возмущения $\xi^j \in \Xi$, причем $\sum_{j=1}^{m} \mu_j = 1$.

На основании исходной вероятностной информации может быть подсчитана вероятностная мера любого подмножества Ξ' в пространстве возмущений:

$$\forall \Xi' \subseteq \Xi \ \ \mu(\Xi') = P\{\xi \in \Xi'\} = \sum_{j:\xi^j \in \Xi'} \mu_j, \tag{23}$$

означающая вероятность P реализации возмущений из этого подмножества.

Для континуальных множеств Ξ нужно прогнозировать плотность вероятности $\rho(\xi): P\{\xi' \in (\xi, \xi + \Delta\xi)\} = \rho(\xi)\Delta\xi$, тогда вместо суммы в (23) нужно вычислять интеграл

$$\mu(\Xi') = \int_{\xi \in \Xi'} \rho(\xi) d\xi. \tag{24}$$

Схема управления остается такой же, как в разд. 2. Оперативные управления u выбираются по тому же алгоритму $\tilde{U}(w, \xi)$, но долговременные управляющие параметры w планируются по-другому. Требование (1) гарантированной допустимости планов w сохраняется.

Вероятностно-гарантирующий подход к планированию, в отличие от оптимизации в среднем, позволяет заранее назначать приемлемую вероятность риска реализации критических возмущений, ухудшающих качество управления по сравнению с прогнозируемой оценкой. Тогда с желаемой вероятностью можно быть уверенным в успехе каждой очередной реализации операции или в успехе уникальной операции.

Формулировка задачи вероятностно-гарантирующего планирования такова: требуется выбрать план $w^{(2)}$, допустимый при любых возмущениях $\xi \in \Xi$ и максимизирующий оценку c критерия $f(w, \xi)$, которая не была бы ухудшена с вероятностью, не меньшей заданного уровня $R \in (0,1]$:

$$c \Rightarrow \max \text{ по } w \in W^0 \text{ и } c: P\{f(w, \xi) \geq c\} \geq R. \tag{25}$$

Фиксированный параметр R в (25) называют надежностью успешного решения.

Для *конечных множеств* Ξ ожидаемых возмущений ξ удобен переборный вариант записи (25) через подмножества $\Xi' \subseteq \Xi$ с достаточной вероятностной мерой $\mu(\Xi') \geq R$:

$$(26) \qquad \max_{w \in W^0} c(w) = c(w^2) = c^2, \text{ где } c(w) = \max_{\Xi' \subseteq \Xi : \mu(\Xi') \geq R} \left[\min_{\xi \in \Xi'} f(w, \xi) \right].$$

Комментарий. В квадратной скобке (26) подсчитывается гарантированная оценка критерия f по подмножеству Ξ'. Затем из всех подмножеств с достаточной вероятностной мерой $\mu(\Xi') = \sum_{\xi^j \in \Xi'} \mu(\xi^j) \geq R$ выбирается наиболее благоприятное, поскольку все такие подмножества равноприемлемы по попаданию в них возмущения. Полученная таким способом оценка качества $c(w)$ максимизируется по гарантированно допустимым управлениям $w \in W^0$.

Предельная тождественность вероятностно-гарантирующего и гарантирующего планов. Когда $R \to 1$, то решение задачи (25), как правило, переходит в решение задачи (2) о наилучшем гарантирующем плане.

Для конечных множеств Ξ это верно, если только $\mu_j > 0$. Для бесконечных – есть достаточные условия. Например, плотность $\rho(\xi)$ конечна, а критерий $f(u, \xi)$ непрерывен по $\xi \in \Xi$ [5].

Вероятностно-гарантированная оценка (25) не выходит за пределы диапазона между осторожной f^0 и оптимистичной f^* оценками:

$$(27) \qquad f^0 \leq c^{(2)}(R) \leq f^* \quad \forall R \in (0, 1].$$

Здесь, также как при оптимизации в среднем, какая-то реализация критерия может оказаться хуже осторожной оценки f^0. Однако вероятность этого события не превышает уровня ($1 - R$), задаваемого самим управляющим в соответствии со своей склонностью к риску.

Для конечных множеств Ξ ожидаемых возмущений ξ^j со строго положительными вероятностями их реализации $\mu_j > 0$ границы диапазона (27) обязательно достигаются

$$(28) \qquad c^{(2)}(R) = f^0 \text{ при } 1 \geq R \geq 1 - \underline{\mu}, \; c^{(2)}(R) = f^* \text{ при } \mu_* \geq R > 0, \text{ где } \underline{\mu} = \min_{j = 1, \ldots, m} \mu_j.$$

Протяженность участков постоянства вероятностно-гарантированной оценки может увеличиваться по сравнению с общими неравенствами из (28).

Для бесконечных множеств Ξ в (27) возможны строгие неравенства, что сопряжено с нарушением упомянутых выше достаточных условий предельной тождественности вероятностно-гарантирующих и гарантирующих решений.

5. Вероятностно-гарантирующее планирование ставок налога и штрафа

Из двух возможных вероятностных подходов к планированию здесь выбран вероятностно-гарантирующий, поскольку долговременные управляющие воздействия w предстоит выбрать для уникальной операции. По этой причине вероятность ухудшения фактических значений критерия по сравнению с его априорной оценкой должна быть не выше некоторого желаемого уровня $(1-R)$, где $R \in (0;1]$ – фиксируемый параметр задачи, названный в разд. 4 надежностью решения.

Чтобы облегчить нерассматриваемый здесь предварительный этап сбора и обработки вероятностной информации о возмущениях и упростить последующую процедуру планирования, вместо трехкомпонентного вектора возмущений из разд. 3 теперь будет использоваться скалярное возмущение $\xi \in [0;1]$, индексирующее поведенческие характеристики налогоплательщиков и налоговой службы.

Семейство рассматриваемых поведенческих характеристик здесь уже не конкретизируется в виде (3) или в виде каких-то других аналитически задаваемых функций. Предполагается только, что их удалось проиндексировать так, что с ростом параметра ξ интенсивность сокрытий λ падает, а эффективности $\omega_{1,2}$ розыска и реализации наказаний растут (рис. 4)

$$\forall \xi' < \xi'' \in [0;1] \ \ \forall w \in W \ \ \lambda(w,\xi') > \lambda(w,\xi''), \ \omega_i(w,\xi') < \omega_i(w,\xi''), \ i = 1,2. \tag{29}$$

Что касается зависимостей λ и $\omega_{1,2}$ от ставок $w_{1,2}$, то их качественные свойства остаются прежними:

$$\lambda \uparrow w_1, \ \lambda \downarrow w_2, \ \omega_i \downarrow w_j \ \ (i,j = 1,2), \tag{30}$$

т.е. интенсивность сокрытий λ монотонно возрастает с увеличением ставки налога w_1 и монотонно уменьшается с ростом ставки штрафа w_2, а интенсивности раскрытий ω_1 и реализации наказаний ω_2 падают и по w_1, и по w_2.

Согласно сделанному предположению (29) рост возмущения ξ означает улучшение ситуации для налоговой службы, выражающееся в уменьшении минимизируемого ею критерия J.

Действительно, благодаря знакоопределенности (9) частных производных критерия $J(\lambda,\omega_{1,2})$ по его исходным аргументам $\lambda,\omega_{1,2}$ итоговая зависимость критерия от возмущения $f(w,\xi)$ как сложная функция окажется монотонно убывающей по ξ:

$$\begin{gathered} f(w,\xi) = J\big(\lambda(w,\xi),\omega_{1,2}(w,\xi)\big) \downarrow \xi, \ \text{ т.е.} \\ \forall \xi' < \xi'' \in [0;1] \ \ \forall w \in W \ \ f_1 = f(w,\xi') > f_2 = f(w,\xi''), \end{gathered} \tag{31}$$

так как в силу предположения (29)

$$\lambda' = \lambda(w,\xi') > \lambda'' = \lambda(w,\xi''),\ \ \omega'_{1,2}(w,\xi') < \omega''_{1,2}(w,\xi''),$$

а

$$f_1 - f_2 = J(\lambda',\omega'_{1,2}) - J(\lambda'',\omega''_{1,2}) = \left[J(\lambda',\omega'_{1,2}) - J(\lambda'',\omega'_{1,2})\right] + \left[J(\lambda'',\omega'_{1,2}) - J(\lambda'',\omega''_{1,2})\right] > 0$$

из-за положительности обеих квадратных скобок, вытекающей из $\partial J/\partial\lambda > 0$ и $\partial J/\partial\omega_i < 0$.

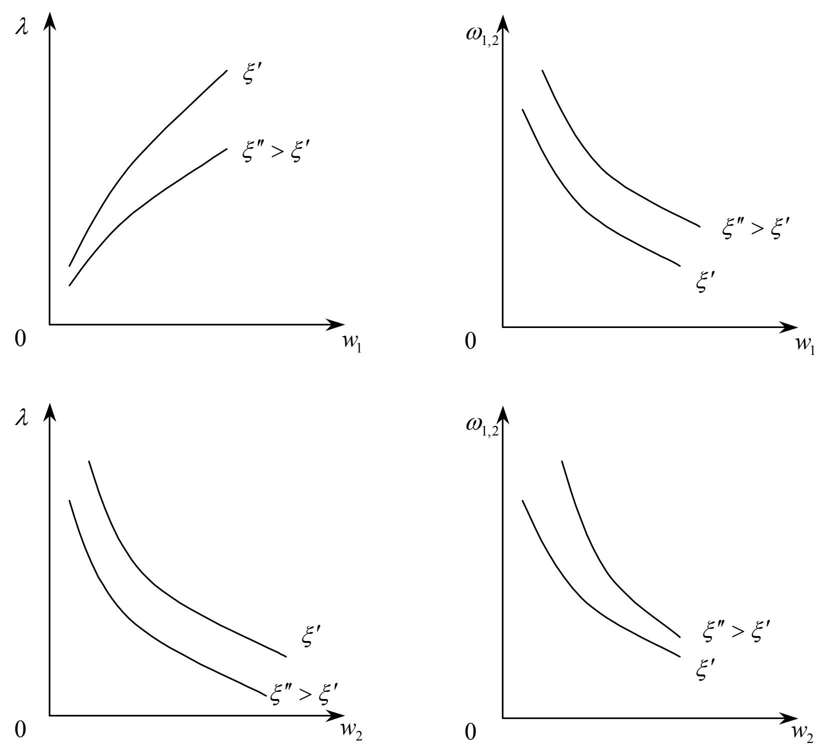

Рис. 4. Характер поведенческих характеристик
при разных значениях скалярного возмущения ξ

Восстановить вероятностные характеристики даже скалярного возмущения по малому числу имеющихся наблюдений весьма трудно. Реально можно надеяться, что объедененными усилиями статистиков и экспертов удастся прогнозировать вероятности μ_j небольшого числа *дискретных значений* ξ_j возмущения:

(32) $$\mu_j = P\{\xi = \xi_j\} > 0,\ \ \xi_j \in [0;1],\ j \in \overline{1,m},\ \sum_{j=1}^{m} \mu_j = 1,\ \xi_j < \xi_{j+1}.$$

Например, $\xi_1 = 0$ − пессимистический прогноз, $\xi_2 = 1/2$ − средний, $\xi_3 = 1$ − оптимистический.

Конечно, подобная дискретизация исходно континуального множества ожидаемых возмущений − это ущемление возможностей возмущения, так как

$$\Xi = \{\xi_1, ..., \xi_m\} \subset [0;1], \tag{33}$$

что может привести к занижению оценок критерия, минимизируемого по w.

Для конечного множества ожидаемых возмущений (33) удобна схема (26) построения вероятностно-гарантирующих планов с использованием подмножеств $\Xi'(R)$ с достаточной вероятностной мерой $\mu(\Xi') \geq R$.

В общей схеме (26) согласно смыслу рассматриваемого сейчас критерия f как суммарного объема нераскрытых и ненаказанных сокрытий нужно заменить максимум на минимум и наоборот:

$$\min_{w \in W}\left[\min_{\Xi':\mu(\Xi')\geq R}\left[\max_{\xi \in \Xi'} f(w,\xi)\right]\right] = c^*(R). \tag{34}$$

Благодаря свойству (31) монотонности критерия по возмущению проблема вероятностно-гарантирующего планирования сводится к детерминированной задаче (16) обычной минимизации, как это удалось сделать для гарантирующего планирования в разд. 2. Разница состоит только в способе задания возмущения ξ, являющегося параметром детерминированной задачи.

В самом деле, на любом подмножестве

$$\Xi' = \left\{\xi_j, j \in J' \subseteq \overline{1,m}\right\}$$

внутренний максимум f по ξ из (34) достигается на самом малом возмущении $\underline{\xi}(\Xi')$

$$\forall w \in W \;\; \forall \Xi' \subseteq \Xi \;\; \max_{\xi \in \Xi'} f(w,\xi) = f\left(w, \underline{\xi}(\Xi')\right), \text{где } \underline{\xi}(\Xi') = \min_{j \in J'} \xi_j, \tag{35}$$

что следует из свойства (31) монотонного роста f при уменьшении ξ.

На том же основании самым лучшим из всех подмножеств Ξ' с достаточной вероятностной мерой

$$\mu(\Xi') = \sum_{j \in J'} \mu_j \geq R$$

будет подмножество с самым большим младшим элементом $\underline{\xi}(\Xi')$, так как

$$\forall \Xi', \Xi'' : \underline{\xi}(\Xi') > \underline{\xi}(\Xi'') \Rightarrow \max_{\xi \in \Xi} f(w,\xi) = f\left(w, \underline{\xi}(\Xi')\right) < f\left(w, \underline{\xi}(\Xi'')\right) = \max_{\xi \in \Xi'} f(w,\xi).$$

Такое подмножество Ξ^* должно содержать все элементы полного множества Ξ, не меньшие младшего $\underline{\xi}(\Xi^*)$:

(36а) $$\Xi^* = \left\{\xi_{j^*}, \xi_{j^*+1}, ..., \xi_m\right\}, \text{ где } \xi_{j^*} = \underline{\xi}\left(\Xi^*\right),$$

а младший элемент не может быть увеличен без нарушения условия достаточной вероятностной меры подмножества, т.е.

(36б) $$\mu\left(\Xi^*\right) = \sum_{j=j^*}^{m} \mu_j \geq R, \text{ но } \sum_{j=j^*+1}^{m} \mu_j < R.$$

Любое другое подмножество Ξ'' с достаточной вероятностной мерой $\mu\left(\Xi''\right) \geq R$ отличается от Ξ^* хотя бы одним элементом $\xi'' \in \Xi''$, $\xi'' \notin \Xi^*$. Причем $\xi'' < \xi_{j^*}$, так как $\forall j \geq j^*$ $\xi_j \in \Xi^*$ согласно (36а), что не выгодно по причине увеличения гарантированной оценки (35):

$$\max_{\xi \in \Xi''} f\left(w, \xi\right) \geq f\left(w, \xi''\right) > f\left(w, \xi_{j^*}\right) = \max_{\xi \in \Xi^*} f\left(w, \xi\right).$$

Таким образом, доказано что задача (34), удовлетворяющая условию (31), эквивалентна детерминированной задаче минимизации:

(37) $$\min_{w \in W} f\left(w, \xi_{j^*}\right) \text{ с } \xi_{j^*} \text{ из (36б).}$$

Чтобы построить оптимальные вероятностно-гарантирующие планы $w^*\left(R\right)$ и найти минимальные оценки критерия $c^*\left(R\right)$ из (34) во всем диапазоне изменения надежности $R \in \left(0;1\right]$, нужно решить серию из m задач (37), придавая параметру ξ_{j^*} все значения от ξ_1 до ξ_m и подсчитывая по (36б) полуинтервалы изменения значения R, в которых сохраняются соответствующие решения:

(38) $$\begin{aligned}
&\min_{w \in W} f\left(w, \xi_1\right) = f\left(w^1, \xi_1\right) = c_1 = f^0 \text{ при } 1 \geq R > 1 - \mu_1, \\
&\min_{w \in W} f\left(w, \xi_2\right) = f\left(w^2, \xi_2\right) = c_2 < c_1 \text{ при } 1 - \mu_1 \geq R > 1 - \mu_1 - \mu_2, \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&\min_{w \in W} f\left(w, \xi_m\right) = f\left(w^m, \xi_m\right) = c_m = f^1 < c_{m-1} \text{ при } \mu_m \geq R > 0.
\end{aligned}$$

На последовательности решения (38) минимальная оценка критерия $c^*\left(R\right)$ ступенчато уменьшается (что хорошо) по мере назначения меньших значений надежности R (что плохо).

Первая самая высокая оценка критерия c_1 совпадает с гарантированной f^0, поскольку согласно (36а) $\Xi^* = \Xi$ при $\xi = \xi_1$. Эта оценка сохранится, пока надежность R не будет уменьшена от единицы настолько, что условие (36б) позволит исключить из Ξ^* возмущение ξ_1.

С этого момента, т.е. с $R \leq 1 - \mu_1$, оценка критерия улучшится до $c_2 < c_1$, так как $\xi_2 > \xi_1$ и выполнено условие (31) о строго монотонном убывании критерия f с ростом ξ.

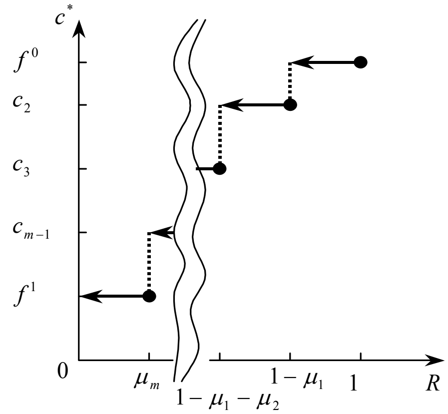

Рис. 5. Ступенчатое улучшение вероятностно-гарантированных оценок критерия c_j с уменьшением надежности решения R

Такое ступенчатое убывание, т.е. улучшение, оценок $c^*(R)$ продолжится пока оптимальное множество Ξ^* достаточной вероятностной меры не удастся сузить до одного самого слабого возмущения ξ_m. Тогда будет достигнута оптимистическая оценка критерия f^1, которая будет справедлива с небольшой надежностью $0 < R < \mu_m$ (если только вероятность реализации самого слабого возмущения ξ_m не близка к единице).

Что касается положений w^j минимумов $f(w, \xi_j)$, т.е. оптимальных ставок налога и штрафа, то для них справедливы те же самые свойства, которые были установлены в разд. 2 для оптимальных гарантирующих планов (с точностью до различий в критических возмущениях).

Процедура построения вероятностно-гарантирующих планов, подобная (38), может быть развита и для *непрерывно распределенного* скалярного возмущения $\xi \in [0;1]$ лишь бы для него сохранилось свойство монотонности (31). В этом случае требуется прежде решить более тонкую статистическую задачу идентификации плотности вероятности $\rho(\xi)$, которая предполагается конечной.

Тогда наилучшее подмножество с достаточной вероятностной мерой $\Xi^*(R)$ вместо (36) будет задаваться интегральным равенством

$$\Xi^*(R) = \left[\underline{\xi}(R), 1\right], \text{ где } \underline{\xi}(R): \int_{\underline{\xi}(R)}^{1} \rho(\xi)\, d\xi = R \in (0;1], \tag{39}$$

а оптимальная вероятностно-гарантирующая программа $w^*(R)$ и наилучшая оценка критерия $c^*(R)$ снова найдутся из решения детерминированной задачи минимизации (37). Только решать ее придется не для конечного числа параметров, как было в (38), а на континууме:

(40) $$\min_{w\in W} f(w,\xi) = f\left(w^*(\xi),\xi\right) = c^*(\xi),\ \xi\in[0;1],$$

где соответствие ξ и R задается интегральным равенством из (39).

Детерминированная задача (40) не содержит плотности вероятности $\rho(\xi)$. Ее просто нужно решить для всех значений параметра $\xi\in[0;1]$, и это уже делалось в разд. 2. Для простоты дальнейших выкладок будем считать, что результирующая зависимость минимальных оценок критерия c^* от параметра ξ аппроксимирована линейной функцией

(41) $$c^* = f^0 - \left(f^0 - f^1\right)\xi,$$

где

$$f^0 = \min_{w\in W}\left[\max_{\xi\in[0;1]} f(w,\xi)\right] = \min_{w\in W} f(w,0) -$$ осторожная оценка,

$$f^1 = \min_{w\in W}\left[\min_{\xi\in[0;1]} f(w,\xi)\right] = \min_{w\in W} f(w,1) -$$ оптимистическая оценка.

Плотность $\rho(\xi)$ нужна для того, чтобы найти соответствие (39) между $\xi = \underline{\xi}$ и R.

Зададим ее в виде семейства аналитически интегрируемых степенных функций, симметричных относительно точки $\xi = 1/2$ и достигающих там максимума (рис. 6):

(42) $$\rho(\xi) = (n+1)2^n \min\{\xi;1-\xi\}^n,\ n\ge 0.$$

При $n=0$ формула (42) дает равномерное распределение, при $0<n<+\infty$ она имитирует нормальное распределение, а при $n\to+\infty$ превращается в сосредоточенное $\rho(\xi)=\delta(\xi-1/2)$.

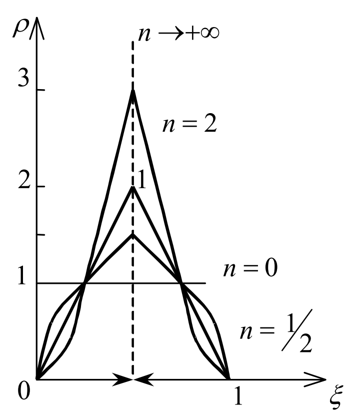

Рис. 6. Семейство степенных функций плотности вероятности (42)

Из интегрального равенства (39) с плотностью (42) находится левая граница $\underline{\xi}(R)$ наилучшего отрезка Ξ^* с достаточной вероятностной мерой (рис. 7):

$$\underline{\xi}(R)=\begin{cases}1-2^{-\frac{n}{n+1}}R^{\frac{1}{n+1}} \text{ при } 0\le R\le\frac{1}{2},\\ 2^{-\frac{n}{n+1}}(1-R)^{\frac{1}{n+1}} \text{ при } \frac{1}{2}\le R\le 1.\end{cases} \tag{43}$$

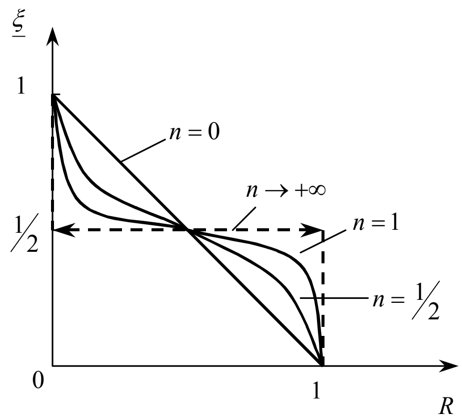

Рис. 7. Функции (43) при разных значениях показателя степени n

Получившиеся функции (43) антисимметричны относительно точки $R=\xi=1/2$. При постоянной плотности ρ, т.е. при $n=0$, это прямая $\underline{\xi}=1-R$. С увеличением n функции становятся зигзагообразными, приближаясь к ступенчатой линии, показанной на рис. 8 жирными штрихами, когда $n\to+\infty$.

В силу линейности аппроксимации (41) функции минимумов $c^*(\xi)$ график $c^*(R)=c^*(\underline{\xi}(R))$ получается зеркальным отражением и сдвигом вверх соответствующего графика для (43)

$$c^*(R)=f^0-(f^0-f^1)\underline{\xi}(R), \tag{44}$$

что показано на рис. 8.

Для рассмотренной задачи выполнены условия из [5] для предельной тождественности вероятностно-гарантирующего решения с $R=1$ и гарантирующего решения, поэтому $c^*(1)=f^0$. При уменьшении надежности R, т.е. при увеличении риска неблагоприятного нарушения оценки критерия, появляется возможность использовать более оптимистические оценки $c^*(R)<f^0$ при $R<1$, монотонно улучшающиеся по мере снижения требований к надежности.

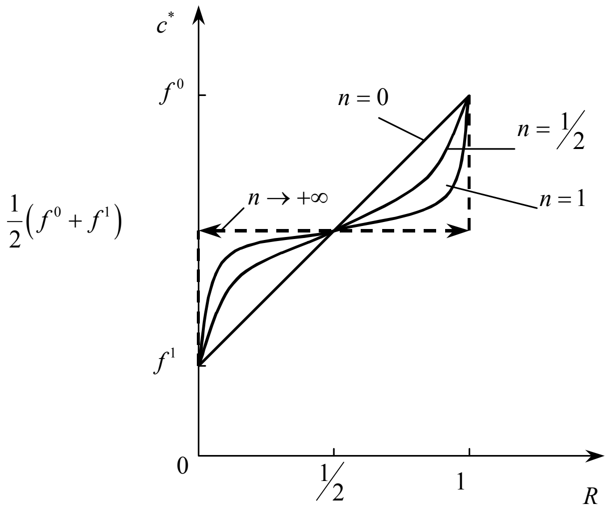

Рис. 8. Улучшение (уменьшение) оценки критерия c^* при снижении требований к ее надежности R

Когда плотность распределения равномерная ($n=0$ в (42)), то согласно (43) левую границу $\underline{\xi}(R)$ отрезка возмущений с достаточной вероятностной мерой удается сдвинуть направо от наихудшей точки $\xi=0$, т.е. в благоприятную сторону, ровно настолько, насколько уменьшена надежность R по сравнению с единицей. Также линейно в соответствии с аппроксимацией (41) откликается на это улучшение оценки критерия.

Если же на границах отрезка $0 \le \xi \le 1$ плотность нулевая ($n>0$ в (42)), то даже при малых уступках в надежности $\varepsilon = 1 - R << 1$ отсекается большой диапазон «плохих» возмущений:

$$0 \le \xi \le 2^{-\frac{n}{n+1}} \varepsilon^{\frac{1}{n+1}}.$$

В пределе, когда $n \to +\infty$, исключается сразу полуинтервал $0 \le \xi < 1/2$. Так что при $n>0$ малые уступки в надежности приводят к большим выигрышам в оценке критерия. Но затем скорость улучшения оценки замедляется. Выбор компромисса между улучшением оценки критерия и ухудшением надежности остается за лицом, принимающим решения, которому присуща большая или меньшая склонность к риску.

6. Основные результаты

- Общие положения гарантирующего и вероятностного планирования при наличии возмущений конкретизированы в приложении к проблеме выбора государством ставок штрафа и налогообложения в условиях неточного прогнозирования поведенческих характеристик налогоплательщиков и налоговой службы.

- Из разнообразных вероятностных подходов в силу уникальности рассматриваемой операции предпочтение отдано вероятностно-гарантирующему, поскольку он обеспечивает желаемую надежность решения. Под этим подразумевается, что вероятность ухудшения реализации критерия по сравнению с прогнозируемой оценкой не превысит заданной величины.
- Установлено, что в рассматриваемой задаче при единичной надежности решения вероятностно-гарантированная оценка критерия совпадает с гарантированной, а с уменьшением надежности она становится все более оптимистичной, причем малые уступки по надежности приводят сначала к большим выигрышам в критерии.
- В задаче доказано наличие седловой точки. С управленческих позиций это означает, что простейший программный способ управления способен обеспечить такую же гарантированную оценку критерия, какая получается для идеального, вообще говоря, нереализуемого управления с точной априорной информацией о возмущениях.
- Благодаря свойствам монотонности критерия по исходным данным, установленным в [4], задачи гарантирующего и вероятностно-гаратирующего планирования удалось свести к статическим задачам детерминированной минимизации, различающимся только способом исчисления критических возмущений.

7. Прикладные выводы

Показано, что с позиции налоговой службы оптимальна самая низкая ставка налога из диапазона оставшихся свобод. Этот вывод справедлив независимо от использования способа планирования, гарантирующего или вероятностно-гарантирующего, при самых общих естественных предположениях о поведенческих характеристиках, лишь бы интенсивность потока новых сокрытий убывала, а эффективность розыска и исполнения наказаний возрастала с уменьшением ставки налога. Однако проблема рационального налогообложения многоаспектна. Здесь нужно принимать в рассмотрение и объем поступлений в государственный бюджет, воздействие на инвестиционный климат и социальные последствия, что выходит за рамки предложенной модели.

Для оптимальной ставки штрафа за сокрытие доходов нет однозначного ответа: она может оказаться и на нижней, и на верхней границах разрешенного диапазона (редко внутри). Ответ зависит от соотношения скоростей уменьшения интенсивности потока сокрытий и падения эффективности работы налоговой службы при увеличении ставки штрафа. Дана классификация поведенческих характеристик, локальная и глобальная, относительно положения точки оптимума по ставке штрафа, но чтобы ей реально воспользоваться нужна довольно тонкая статистика. При отсутствии таковой удается определить только направление улучшающих изменений ставки штрафа.

* *
*

СПИСОК ЛИТЕРАТУРЫ

1. *Гермейер Ю.Б.* Введение в теорию исследования операций. М.: Наука, 1971.

2. *Гермейер Ю.Б.* Игры с непротивоположными интересами. М.: Наука, 1976.

3. *Ермольев Ю.М., Ястремский А.И.* Стохастические модели и методы в экономическом планировании. М.: Наука, 1979.

4. *Закревский А.В., Токарев В.В.* Оптимизация распределения ресурсов между раскрытием и наказанием уклонений от налогов // Экономический журнал ВШЭ. 2003. Т. 7. № 4.

5. *Кича И.В., Токарев В.В.* Вероятностное и гарантирующее управление. III. Предельная тождественность // Автоматика и телемеханика. 1994. № 10. С. 143–150.

6. *Токарев В.В.* Вероятностное и гарантирующее управление в экономике // Управление экономикой переходного периода. Вып. 2 / Под ред. В.В. Макарова. М.: Наука, 1998.

7. *Токарев В.В.* Гарантирующий договор и оперативная компенсация сбоев в сырьевых поставках. I-III // Автоматика и телемеханика. 1992. № 10 С. 120–126; № 11. С. 117–126; № 12. С. 81–87.

8. *Токарев В.В.* Планируемые договоры и оперативные сделки // Управление экономикой переходного периода. Вып. 3 / Под ред. В.В. Макарова. М.: Наука, 1998.