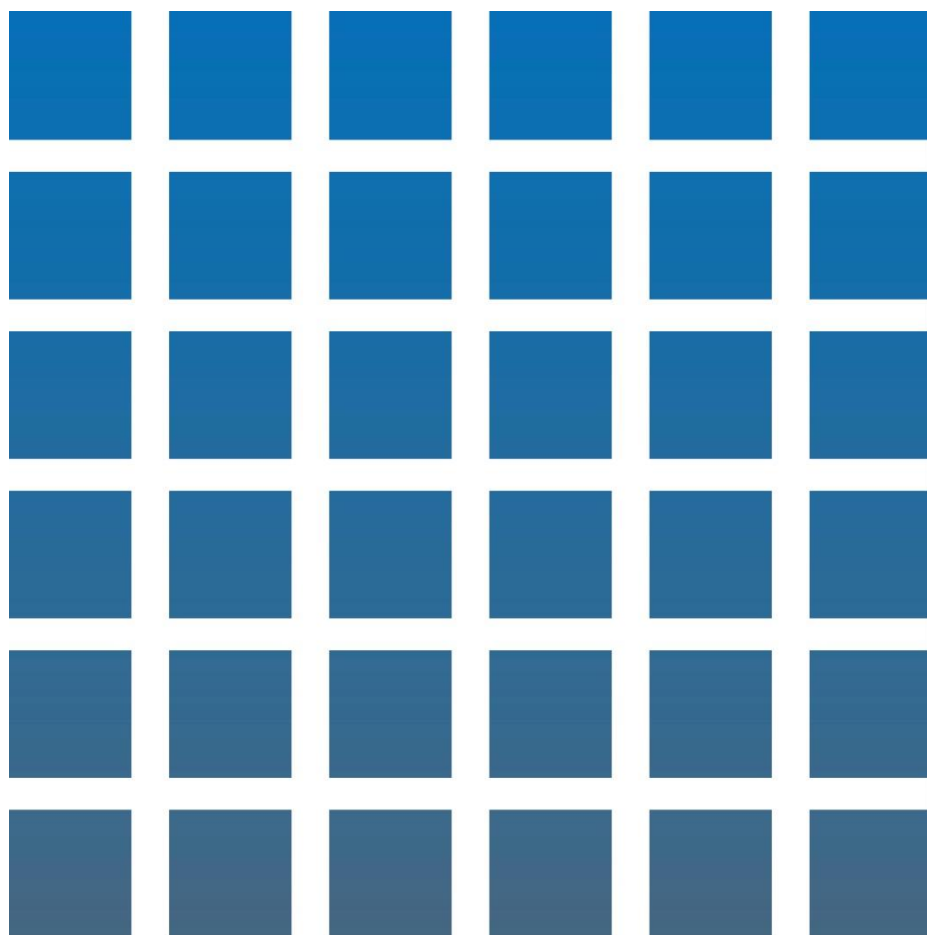


Modellierung der Zusammenhänge zwischen ökonomischen Indikatoren, Sollerträgen und Einnahmen der direkten Bundessteuer auf Basis desaggregierter Daten

Schlussbericht z.H. der Eidgenössischen Finanzverwaltung EFV & der Eidgenössischen Steuerverwaltung ESTV

06. Mai 2016



Herausgeber

BAK Basel Economics AG

Projektleitung

Martin Eichler, T +41 61 279 97 14
martin.eichler@bakbasel.com

Michael Grass, T +41 61 279 97 23
michael.grass@bakbasel.com

Redaktion

Mark Emmenegger, T +41 61 279 97 29
mark.emmenegger@bakbasel.com

Klaus Jank, T +41 61 279 97 24
klaus.jank@bakbasel.com

Adresse

BAK Basel Economics AG
Güterstrasse 82
CH-4053 Basel
T +41 61 279 97 00
F +41 61 279 97 28
info@bakbasel.com
www.bakbasel.com

Executive Summary

Dieses Gutachten im Auftrag der Eidgenössischen Finanzverwaltung EFV und der Eidgenössischen Steuerverwaltung ESTV soll einen Beitrag zur Überprüfung der aktuellen Schätzmethode für die direkte Bundessteuer leisten sowie gegebenenfalls Ansatzpunkte für eine zukünftige Weiterentwicklung der Methoden liefern. Konkret besteht das Ziel der Studie darin, dem bisherigen Schätzmodell zugrundeliegende Zusammenhänge auf der Basis von desaggregierten Daten in einer ex-post Analyse zu untersuchen, um herauszufinden, ob in einer desaggregierten Modellierung Potenzial für eine Verbesserung der Prognosequalität liegt.¹ Dabei geht es um die folgenden Zusammenhänge: erstens, um den Zusammenhang zwischen makroökonomischen Indikatoren und den Sollerträgen der Einkommens- bzw. Gewinnsteuern; zweitens, um den (zeitlichen) Zusammenhang zwischen den Sollerträgen und den Steuereinnahmen. Diese beiden Zusammenhänge werden separat analysiert. Bezüglich der Desaggregation liegt der primäre Fokus auf einer regionalen bzw. kantonalen Desaggregation; beim Zusammenhang zwischen Indikatoren und Sollerträgen wird eine zusätzliche Desaggregation nach Einkommensklassen (bei der Einkommenssteuer) bzw. nach Steuerstatus (bei der Gewinnsteuer) untersucht.

Zusammenhang zwischen Indikatoren und Sollerträgen

Die Vorgehensweise bei der Analyse des ersten Zusammenhangs besteht darin, von einer einfachen Modellierung (Benchmarks) auszugehen und in einer schrittweisen Sophistizierung der Schätzungen zu prüfen, inwiefern damit eine Verbesserung der ex-post Prognosequalität erreicht werden kann. In der einfachen Modellierung werden die Zusammenhänge zwischen den Sollerträgen und den im aktuellen Schätzmodell verwendeten Indikatoren – Primäreinkommen bei der Einkommenssteuer, BIP bei der Gewinnsteuer – auf der Basis aggregierter/nationaler Daten modelliert. Relativ dazu wird auf zweifache Weise eine Sophistizierung vollzogen: zum einen wird der Zusammenhang auf der Basis desaggregierter Daten modelliert (kantonale Daten und zusätzliche Berücksichtigung von Einkommensklassen bzw. Steuerstatus), zum anderen werden im Rahmen der Desaggregation auch alternative Indikatoren zu den im aktuellen Schätzmodell verwendeten in Betracht gezogen. Methodologisch liegt der Untersuchung ein Regressionsmodell zugrunde, welches für eine grosse Anzahl von Kombinationen aus Sollertragsvariablen und Indikatoren sowie verschiedene Lag-Strukturen geschätzt wurde.

Die Hauptergebnisse zur Einkommenssteuer lauten folgendermassen:

Eine kantonale Desaggregation auf der Basis des theoretisch sinnvollsten Indikators – der kantonalen pendlerbereinigten Lohnsummen – reicht nicht aus, um die ex-post Prognosegüte zu erhöhen. Erst bei der Verwendung eines kantonsspezifischen Indikatoren-Sets aus mehreren ökonomischen Grössen (Primäreinkommen CH, Lohnsumme CH und Kantone, Nominales BIP CH und Kantone sowie Bruttowertschöpfung CH und Kantone) kann die Schätzgenauigkeit verbessert werden. Dies deutet darauf hin, dass die Einkommenssteuer-Entwicklung in den Kantonen unterschiedlich verläuft und von abweichenden ökonomischen Erklärungsfaktoren bestimmt wird. Durch

¹ Eine Analyse der ex-ante Prognosequalität, wie sie für die Weiterentwicklung des Prognosemodells zusätzlich notwendig ist, war explizit nicht Teil der hier vorliegenden Analyse.

die Verwendung von weiteren ökonomischen Indikatoren (z.B. Lohnsumme Finanzsektor in einigen Kantonen) kann die ex-post Prognosequalität noch weiter verbessert werden, allerdings steigt dadurch auch die Gefahr von zufälligen statistischen Korrelationen aufgrund der recht kurzen Beobachtungsperiode. Ferner muss man sich bewusst sein, dass diese Ergebnisse auf einer ex-post Untersuchung beruhen und nicht ohne weiteres in den ex-ante Kontext übertragen werden können. Insgesamt erachten wir die Desaggregation nach Kantonen als eine prüfenswerte Erweiterung der bisherigen Schätzmethoden für die Einkommenssteuer. Bevor aber eine Erweiterung der bisherigen Schätzmethoden in Betracht gezogen würde, müssten in einem nächsten Schritt die besten desaggregierten Schätzmethoden und die ihnen zugrundeliegenden Indikatoren noch auf ihre ex-ante Prognosequalität untersucht werden.

Die zusätzliche Desaggregation nach Einkommensklassen scheint dagegen keine weitere Verbesserung zu bringen. Sie ist daher aufgrund des steigenden Aufwandes zumindest bei der gegenwärtigen Datenlage nicht zu empfehlen.

Die Hauptergebnisse zur Gewinnsteuer können wie folgt zusammengefasst werden:

Bei den Gewinnsteuererträgen ist es schwierig, die ex-post Prognosequalität alleine auf der Basis von einer kantonalen Desaggregation zu verbessern. Eine deutliche Verbesserung der Prognosequalität wird erst erreicht, wenn das Regressionsmodell um einen finanzplatzbezogenen Interaktionsterm für die beiden Finanzkrisen 2001 und 2008 ergänzt wird. Eine etwaige Verwendung dieser Methode im Rahmen von ex-ante Prognosen würde jedoch bedingen, dass frühzeitig genug antizipiert werden kann, ob es zu einer Finanzkrise kommt oder nicht.

Die Unterteilung nach Steuerstatus der Unternehmen als zusätzliche Desaggregationsstufe verbessert die ex-post Prognosequalität bei allen getesteten Schätzmethoden. Man muss aber berücksichtigen, dass der Zeitraum für die Beurteilung der ex-post Prognosequalität aufgrund der Datenlage bei den Schätzungen mit einer zusätzlichen Desaggregation nach Steuerstatus kleiner ist als bei den Schätzungen ohne eine solche Desaggregation, was die Ergebnisse weniger belastbar macht. Die Ergebnisse deuten aber zumindest darauf hin, dass im Steuerstatus für die Prognose relevante Informationen enthalten sein könnten. Denn den Steuerstatus kann man als Proxy für Unternehmensmerkmale wie Branchenzugehörigkeit, Internationalität der Gesellschaft, Unternehmensgrösse etc. interpretieren, über welche gegenwärtig keine desaggregierten Daten vorliegen. Wir erachten es deshalb auch angesichts der wahrscheinlichen Abschaffung des Steuerstatus im Zuge der Unternehmenssteuerreform III als prüfenswert, diese Unternehmensmerkmale in Zukunft systematisch zu erheben und in einigen Jahren auf das mit ihnen möglicherweise verbundene Verbesserungspotenzial zu untersuchen.

Zeitlicher Zusammenhang zwischen Sollerträgen und Steuereinnahmen

Bei der Untersuchung des zeitlichen Zusammenhangs zwischen den Sollerträgen und den Steuereinnahmen geht es um zeitliche Verteilungsmuster, das heisst, wie sich die Steuererträge (Sollerträge) eines bestimmten Bemessungsjahres (t_0) auf die Steuereinnahmen des entsprechenden Jahres (t_0) und der nachfolgenden Jahre (t_1 , t_2 , t_3 ...) aufteilen. In einem ersten Schritt werden mit deskriptiven statistischen Methoden interkantonale und intertemporale Unterschiede in den zeitlichen Verteilungsmustern untersucht. Mit intertemporalen Unterschieden in den zeitlichen Vertei-

lungsmustern sind Trends im Ablieferungsverhalten von natürlichen und juristischen Personen gemeint. In einem zweiten Schritt werden die Steuereinnahmen auf der Basis von (tatsächlichen) Sollerträgen aus der Steuerstatistik und zeitlichen Verteilungsmustern modelliert. Dabei wird insbesondere überprüft, ob die ex-post Prognosequalität der Steuereinnahmen durch eine kantonale desaggregierte Modellierung verbessert werden kann. Aufgrund der zuvor identifizierten Trends in den Verteilungsmustern wird in der Modellierung exemplarisch mit zwei verschiedenen Mustern gearbeitet: einem Muster, welches auf der Basis weiter zurückreichender Zeitreihen berechnet wurde und einem Muster, welches auf einem aktuelleren Zeitraum beruht.

Die Hauptergebnisse zu diesem zweiten Zusammenhang lauten folgendermassen: Die deskriptive Analyse hat zutage gefördert, dass es beträchtliche kantonale und intertemporale Variationen in den Verteilungsmustern gibt. In der anschliessenden Modellierung konnte zum einen gezeigt werden, dass eine Modellierung des Zusammenhangs auf der Basis kantonalen Daten die ex-post Prognosegüte leicht verbessert. Da die Verbesserung aber verhältnismässig gering ausfällt, scheint bei diesem Zusammenhang eine kantonale Desaggregation zur Optimierung des bisherigen Schätzverfahrens eher wenig geeignet. Zum anderen hat sich gezeigt, dass es im Vergleich zur Frage, ob man die Kantonsvarianz berücksichtigen sollte oder nicht, eigentlich eine wichtigere Rolle spielt, welchen Zeitraum man zur Berechnung der durchschnittlichen Verteilungsmuster wählt. Die ESTV trägt diesem Umstand bereits insofern Rechnung, als die im Schätzmodell verwendeten Verteilungsmuster laufend überprüft und angepasst werden. Angesichts der erheblichen Verschiebung der Verteilungsmuster über die Zeit könnte hier jedoch noch weiteres Verbesserungspotential liegen, insbesondere wenn es gelingen würde, die Ursachen für die Veränderungen in die Modellierung mit einzubeziehen (z.B. konjunkturelle Einflüsse, regulatorische Veränderungen).

Verbesserung der Datenlage

Über diese technischen Zusammenhänge hinaus ist in unserer Arbeit mehrfach deutlich geworden, dass die Datenlage für die Prognose der direkten Bundessteuer auf Seiten der Steuerdaten (Erträge und Einnahmen) spärlich ist. Es erscheint uns deshalb für zukünftige Verbesserungen der Prognosegüte sinnvoll, wenn die Datenlage systematisch verbessert würde. Dies würde unter anderem Folgendes beinhalten: Erstens eine Erhebung von präzisen Daten zu den Verlustvorträgen von Unternehmen. Zweitens nach Unternehmensmerkmalen wie Branche, Grösse und Internationalität desaggregierte Gewinnsteuerdaten. Dies insbesondere deshalb, weil die Studie zeigt, dass mit der Berücksichtigung des Steuerstatus, welcher als Proxy für solche Merkmale gesehen werden kann, die ex-post Prognosequalität verbessert werden kann. Drittens nach natürlichen und juristischen Personen getrennte Daten zum Ablieferungsverhalten. Und viertens ist es aus akademischer wie aus budgetärer Sicht grundsätzlich anzustreben, dass alle (weiteren) nicht-sensiblen Daten, welche bisher nur den kantonalen Steuerbehörden vorliegen, schweizweit harmonisiert und so zeitnah wie möglich zugänglich gemacht werden.

Condensé

La présente expertise établie à la demande de l'Administration fédérale des finances (AFF) et de l'Administration fédérale des contributions (AFC) vise à contribuer à l'examen des actuelles méthodes d'estimation du produit de l'impôt fédéral direct, et, le cas échéant, à fournir des suggestions pour le développement de ces méthodes. Elle a concrètement pour but d'examiner, à partir de données désagrégées et dans le cadre d'une analyse ex post, les relations sur lesquelles repose le modèle d'estimation actuel, afin d'établir si une modélisation désagrégée recèlerait un potentiel d'amélioration pour la qualité des prévisions. Les relations en question sont, d'une part, le rapport entre les indicateurs macroéconomiques et le produit théorique des impôts sur le revenu et sur le bénéfice et, d'autre part, le lien (temporel) entre les produits théoriques et les recettes fiscales. Ces deux relations sont analysées séparément. S'agissant de la désagrégation, la priorité est donnée à une approche régionale, c'est-à-dire cantonale; s'agissant de la relation entre indicateurs et produits théoriques, l'examen porte, en outre, sur une désagrégation par classe de revenus (pour l'impôt sur le revenu) ou par statut fiscal (pour l'impôt sur les bénéfices).

Rapport entre indicateurs et produits théoriques

La méthode employée pour analyser la première relation consiste à partir d'une modélisation simple (benchmarks) et à vérifier, en augmentant progressivement la complexité des estimations, dans quelle mesure cela permet d'améliorer la qualité des prévisions ex post. L'approche simple consiste à modéliser les relations entre les produits théoriques et les indicateurs employés dans le modèle d'estimation actuel (revenus primaires pour l'impôt sur le revenu, PIB pour l'impôt sur le bénéfice) sur la base de données agrégées / nationales. A partir de là, on procède à une double sophistication: d'un côté, on modélise la relation sur la base de données désagrégées (données cantonales et prise en compte, en plus, de la classe de revenus ou du statut fiscal) et, de l'autre, on prend en compte, dans le cadre de la désagrégation, des indicateurs différents de ceux employés dans le modèle d'estimation actuel. Sur le plan méthodologique, l'analyse repose sur un modèle de régression estimé pour un grand nombre de combinaisons de variables relatives aux produits théoriques et d'indicateurs, de même que pour plusieurs structures contextuelles.

Les principales conclusions concernant l'impôt sur le revenu sont les suivantes:

Une désagrégation par canton effectuée sur la base de l'indicateur le plus approprié théoriquement (la masse salariale du canton corrigée des navetteurs) ne suffit pas à augmenter la précision des prévisions ex post. Cette précision n'augmente que si l'on utilise un ensemble d'indicateurs spécifique au canton composé de plusieurs grandeurs économiques (revenus primaires Confédération, masse salariale Confédération et cantons, PIB nominal Confédération et cantons et valeur ajoutée brute Confédération et cantons), ce qui donne à penser que l'évolution de l'impôt sur le revenu varie selon le canton et dépend de facteurs économiques différents. L'utilisation d'indicateurs économiques supplémentaires (masse salariale du secteur financier dans certains cantons, par ex.) permet d'améliorer encore plus la qualité des prévisions ex post, mais elle fait augmenter le risque de corrélations statistiques fortuites en raison de la brièveté de la période observée. Il faut par ailleurs garder à l'esprit que ces

résultats reposent sur un examen ex post et ne peuvent pas être transposés comme cela dans un contexte ex ante. Dans l'ensemble, nous considérons la désagrégation par canton comme un élargissement digne d'être examiné des méthodes d'estimation actuelles pour l'impôt sur le revenu. Avant toutefois d'envisager un tel élargissement, il faudrait commencer par examiner les meilleures méthodes d'estimation fondées sur une approche désagrégée de même que leurs indicateurs sous-jacents quant à la qualité de leurs prévisions ex ante.

La désagrégation par classe de revenu ne semble par contre apporter aucune amélioration. Nous ne la recommandons donc pas en l'état actuel des données, vu la complexité croissante des tâches.

S'agissant de l'impôt sur le bénéfice, les principales conclusions sont les suivantes:

Dans le cas du produit de l'impôt sur le bénéfice, il est difficile d'améliorer la qualité des prévisions ex post sur la seule base d'une désagrégation par canton. On n'obtient une amélioration sensible qu'en complétant le modèle de régression avec un terme d'interaction se rapportant à la place financière pour les deux crises financières de 2001 et 2008. L'éventuel recours à cette méthode dans le cadre de prévisions ex ante supposerait toutefois que l'on puisse anticiper suffisamment la survenue ou non d'une crise financière.

La subdivision par statut fiscal des entreprises comme échelon de désagrégation supplémentaire améliore la qualité des prévisions ex post avec toutes les méthodes examinées. Il faut toutefois tenir compte du fait que dans ce cas, la période d'évaluation de cette qualité est moins longue, en raison de l'état des données, que pour les évaluations sans désagrégation supplémentaire, ce qui rend les résultats moins fiables. Quoi qu'il en soit, les résultats indiquent que le statut fiscal peut contenir des informations pertinentes pour les prévisions. En effet, ce statut peut s'interpréter comme le représentant de caractéristiques de l'entreprise telles que son appartenance à un secteur, son caractère international, sa taille, etc., pour lesquelles il n'existe pas de données désagrégées pour le moment. Nous considérons par conséquent, eu égard également à l'abolition probable du statut fiscal dans la foulée de la réforme sur l'imposition des entreprises III, qu'il vaut la peine d'envisager le recueil systématique de ces caractéristiques à l'avenir, et d'examiner dans quelques années le potentiel d'amélioration qu'elles peuvent receler.

Lien temporel entre produits théoriques et recettes fiscales

L'examen de ce lien porte sur des modèles de répartition temporels, c'est-à-dire sur la manière dont les recettes fiscales (produits théoriques) d'une année t_0 se répartiront sur les recettes fiscales de cette année t_0 et des années suivantes t_1 , t_2 , t_3 ... Dans un premier temps, on examine au moyen de méthodes statistiques descriptives les écarts intercantonaux et intertemporels dans les modèles de répartition temporels. Par écarts intertemporels dans les modèles de répartition temporels, on entend les tendances des personnes physiques ou morales en matière d'acquiescement de l'impôt. Dans un deuxième temps, on modélise les recettes fiscales sur la base de produits théoriques (effectifs) issus des statistiques fiscales et de modèles de répartition temporels. On vérifie en particulier s'il est possible d'améliorer la qualité des prévisions ex post concernant les recettes fiscales au moyen d'une modélisation désagrégée. Sur la base des tendances identifiées préalablement dans les modèles de

répartition, on emploie à titre d'exemple deux modèles différents pour la modélisation, l'un calculé à partir de séries chronologiques remontant plus loin dans le passé, l'autre reposant sur une période plus récente.

Les principaux résultats concernant ce deuxième lien sont les suivants: L'analyse descriptive a révélé qu'il existe des écarts cantonaux et intertemporels considérables dans les modèles de répartition. Lors de la modélisation qui a suivi, on a pu établir d'une part qu'une modélisation de la relation sur la base de données cantonales améliore légèrement la qualité des prévisions ex post. Comme cette amélioration est relativement faible, il paraît cependant peu judicieux de recourir à une désagrégation par canton pour améliorer la méthode d'évaluation employée jusque-là. D'autre part, il est apparu que la question de savoir s'il faut prendre en compte les variations cantonales est moins importante qu'une sélection soignée de la période de calcul des modèles de répartition moyens. L'AFC en tient d'ores et déjà compte dans la mesure où elle examine constamment, pour les adapter, les modèles de répartition utilisés dans le modèle d'estimation. Compte tenu du glissement considérable des modèles de répartition au fil du temps, il se pourrait que le potentiel d'amélioration correspondant soit encore plus grand, surtout si l'on parvient à intégrer les causes du changement dans la modélisation (facteurs conjoncturels, évolution de la réglementation, etc.).

Amélioration de l'état des données

Au-delà de ces relations techniques, notre travail nous a fait prendre conscience à plusieurs reprises du caractère lacunaire des données fiscales (produits et recettes) sur lesquelles fonder une estimation du produit de l'impôt fédéral direct. Il nous paraît donc judicieux, pour améliorer la qualité des prévisions, d'améliorer systématiquement l'état des données. Cela supposerait notamment les démarches suivantes: premièrement, recueillir des données précises sur les pertes reportées des entreprises; deuxièmement, recueillir des données concernant l'impôt sur le bénéfice désagrégées en fonction de caractéristiques de l'entreprise telles que secteur, taille et caractère international, ce qui nous paraît d'autant plus important que l'étude montre que la prise en compte du statut fiscal, sorte de substitut de ce genre de caractéristiques, permet d'améliorer la qualité des prévisions ex post; troisièmement, recueillir des données séparées sur les tendances des personnes physiques et morales en matière d'acquiescement de l'impôt et, quatrièmement, viser, eu égard à des considérations tant académiques que budgétaires, l'harmonisation à l'échelle de la Suisse et la mise à disposition aussi rapide que possible de toutes les (autres) données non sensibles jusqu'ici réservées aux autorités fiscales cantonales.

Inhalt

Executive Summary	3
Condensé	6
1 Einleitung	11
2 Zusammenhang Indikatoren und Sollerträge	12
2.1 Untersuchungsansatz.....	12
2.2 Verwendete Daten	13
2.2.1 Steuererträge	13
2.2.2 Indikatoren aus den BAKBASEL Datenbanken.....	14
2.3 Methodologische Vorgehensweise	15
2.3.1 Regressionsmodell	15
2.3.2 Schätzung der Benchmarks.....	16
2.3.3 Schätzung des Modells für eine Vielzahl von Steuerertrags-Indikatoren- Kombinationen	17
2.3.4 Auswahl geeigneter Indikatoren-Sets.....	17
2.3.5 Backtest der gewählten Indikatoren-Sets.....	20
2.4 Ergebnisse für die Einkommenssteuer	21
2.4.1 Überblick	21
2.4.1 Schätzmethode S1	22
2.4.2 Schätzmethode S2	23
2.4.3 Schätzmethode S3	24
2.4.4 Schätzmethoden S4a und S4b.....	25
2.4.1 Schätzmethode S5	27
2.5 Ergebnisse für die Gewinnsteuer.....	28
2.5.1 Überblick	28
2.5.2 Schätzmethode S1	30
2.5.3 Schätzmethode S2	30
2.5.4 Schätzmethode S3	31
2.5.5 Schätzmethoden S4a und S4b.....	31
2.5.6 Schätzmethode S5	34
2.6 Fazit	34
3 Zusammenhang Sollerträge und Steuereinnahmen	37
3.1 Untersuchungsansatz.....	37
3.2 Verwendete Daten	37
3.3 Unterschiede in den Verteilungsmustern zwischen Kantonen.....	38
3.4 Unterschiede in den Verteilungsmustern über die Zeit.....	40
3.5 Methodologische Vorgehensweise bei der Modellierung	42
3.6 Ergebnisse der Modellierung	44
3.7 Fazit	46
4 Schlussfolgerungen und Empfehlungen	47
5 Literatur	51

Tabellenverzeichnis

Tab. 2-1	Vergleich der ex-post Prognosequalität von verschiedenen Benchmarks und Schätzmethoden für die Sollerträge der Einkommenssteuer	22
Tab. 2-2	Vergleich der ex-post Prognosequalität von verschiedenen Benchmarks und Schätzmethoden für die Sollerträge der Gewinnsteuer.....	29
Tab. 3-1	Vergleich der ex-post Prognosequalität von verschiedenen Benchmarks und Schätzmethoden für die Steuereinnahmen	44

Abbildungsverzeichnis

Abb. 2-1	S1: Abweichung des geschätzten Steuerertrags pro Jahr	23
Abb. 2-2	S2: Abweichung des geschätzten Steuerertrags pro Jahr	24
Abb. 2-3	S3: Abweichung des geschätzten Steuerertrags pro Jahr	25
Abb. 2-4	S4b: Abweichung des geschätzten Steuerertrags.....	26
Abb. 2-5	S5: Abweichung des geschätzten Steuerertrags.....	27
Abb. 2-6	S2: Abweichung des geschätzten Steuerertrags pro Jahr	30
Abb. 2-7	S2 mit Berücksichtigung des Steuerstatus: Abweichung des geschätzten Steuerertrags pro Jahr	31
Abb. 2-8	S4b: Abweichung des geschätzten Steuerertrags pro Jahr	33
Abb. 2-9	S4b mit Berücksichtigung des Steuerstatus: Abweichung des geschätzten Steuerertrags pro Jahr	33
Abb. 3-1	Kantonale Unterschiede in den Verteilungsmustern: Durchschnittliche (1995 – 2011) pro Jahr beglichene Steuerschuld in Prozent	39
Abb. 3-2	Kantonale Unterschiede in den Verteilungsmustern bei den 5 grössten Steuerzahler-Kantonen: Durchschnittliche (1995 – 2011) pro Jahr beglichene Steuerschuld in Prozent	39
Abb. 3-3	Zeitliche Veränderung in den Verteilungsmustern: Prozentualer Anteil der beglichene gesamtschweizerischen Steuerschuld in den Jahren t_0 bis t_3 über die Zeit	40
Abb. 3-4	Zeitliche Veränderungen in den Verteilungsmustern: Entwicklung Ablieferungsanteils von t_0 über die Zeit in ausgewählten Kantonen und der Schweiz	41
Abb. 3-5	Abweichung der geschätzten Steuereinnahmen pro Jahr	46

1 Einleitung

Zwischen 2010 und 2014 expandierte die Schweizer Wirtschaft trotz vielfältiger Belastungen wie zum Beispiel der Frankenstärke, gleichzeitig stagnierten aber die Steuereinnahmen aus der direkten Bundessteuer, welche neben der Mehrwertsteuer die wichtigste Einnahmequelle des Bundes darstellen. Da mit einer positiven Wirtschaftsentwicklung üblicherweise steigende Steuereinnahmen einhergehen, wurde bei der Budgetierung für die Einkommens- und Gewinnsteuern mit höheren Einnahmen gerechnet als tatsächlich realisiert werden konnten. Der Bund zieht daraus unter anderem die Konsequenz, das bestehende Schätzmodell für die direkte Bundessteuer zu überprüfen.²

Dieses Gutachten im Auftrag der Eidgenössischen Finanzverwaltung EFV und der Eidgenössischen Steuerverwaltung ESTV soll einen Beitrag zur Überprüfung der aktuellen Schätzmethoden leisten sowie gegebenenfalls Ansatzpunkte für eine zukünftige Weiterentwicklung der Methoden liefern. Genauer formuliert besteht das Ziel der Studie darin, dem bisherigen Schätzmodell zugrundeliegende Zusammenhänge auf der Basis von desaggregierten Daten in einer ex-post Analyse zu untersuchen, um herauszufinden, ob in einer desaggregierten Modellierung Potenzial für eine Verbesserung der Prognosequalität liegt.³ Dabei geht es um die folgenden Zusammenhänge: erstens, um den Zusammenhang zwischen makroökonomischen Indikatoren und den Sollerträgen der Einkommens- bzw. Gewinnsteuern; zweitens, um den (zeitlichen) Zusammenhang zwischen den Sollerträgen und den Steuereinnahmen. Diese beiden Zusammenhänge werden separat analysiert. Bezüglich der Desaggregation liegt der primäre Fokus auf einer regionalen bzw. kantonalen Desaggregation; beim ersten Zusammenhang wird eine zusätzliche Desaggregation nach Einkommensklassen (bei der Einkommenssteuer) bzw. nach Steuerstatus (bei der Gewinnsteuer) untersucht.

Die Studie ist folgendermassen aufgebaut: Im Kapitel 2 wird der erste Zusammenhang untersucht, d.h. der Zusammenhang zwischen makroökonomischen Indikatoren und Sollerträgen aus der Einkommens- bzw. Gewinnsteuer. Unsere Untersuchungsstrategie besteht hier darin, von einer relativ einfachen aggregierten Modellierung (Benchmarks) auszugehen und in einer schrittweisen Sophistizierung (Sollertragschätzungen auf der Basis von Desaggregation und alternativer Indikatoren) zu prüfen, ob und in welchem Mass damit die ex-post Prognosequalität verbessert werden kann. Im Kapitel 3 folgt die Analyse des zweiten Zusammenhangs, d.h. des zeitlichen Verteilungsmusters zwischen den Sollerträgen und den Steuereinnahmen. Unsere Analyse umfasst hier zunächst eine deskriptive Untersuchung von interkantonalen und intertemporalen Unterschieden in den zeitlichen Verteilungsmustern. Danach geht es darum zu untersuchen, welchen Vorteil eine regionale Desaggregation unter der Annahme verschiedener Verteilungsmuster für die Modellierung des Zusammenhangs bringt. Im abschliessenden Kapitel 4 werden die Ergebnisse zusammengefasst und Empfehlungen für die Weiterentwicklung bzw. Ergänzung des Prognosemodells abgeleitet.

² Das gegenwärtige Schätzmodell ist in ESTV (2015) beschrieben.

³ Das Ziel der Studie besteht aber nicht darin, die Prognosegüte des bisherigen Schätzmodells für die direkte Bundessteuer zu evaluieren und auch nicht darin, das bisherige Schätzmodell bereits konkret weiterzuentwickeln.

2 Zusammenhang Indikatoren und Sollerträge

In diesem Kapitel wird der statistische Zusammenhang zwischen ausgewählten ökonomischen Variablen der Regionen- bzw. Branchendatenbanken von BAKBASEL und dem Sollertrag der natürlichen bzw. der juristischen Personen untersucht. Es wird insbesondere überprüft, ob mithilfe von desaggregierten Daten und Schätzungen eine Verbesserung der ex-post Prognosequalität erzielt werden kann. Dieses Kapitel ist folgendermassen aufgebaut: Zuerst wird der Untersuchungsansatz eingeführt; danach werden die verwendeten Daten kurz erläutert, bevor ausführlich auf die Methoden eingegangen wird; anschliessend werden die Ergebnisse vorgestellt und zum Schluss ein Fazit zum Zusammenhang zwischen makroökonomischen Grössen und den Sollerträgen gezogen.

2.1 Untersuchungsansatz

Unsere Vorgehensweise bei der Untersuchung des Zusammenhangs zwischen makroökonomischen Indikatoren und den Sollerträgen ist für die Einkommens- und Gewinnsteuern nahezu identisch. Der Grundgedanke besteht darin, von einer einfachen Modellierung der Zusammenhänge (den sogenannten Benchmarks) auszugehen und in einer schrittweisen Sophistizierung der Schätzungen zu prüfen, inwiefern damit eine Verbesserung der ex-post Prognosequalität erreicht werden kann. In der einfachen Modellierung werden die Zusammenhänge zwischen den Sollerträgen und den im aktuellen Schätzmodell verwendeten Indikatoren – Primäreinkommen bei der Einkommenssteuer, BIP bei der Gewinnsteuer – auf der Basis aggregierter/nationaler Daten modelliert. Dabei ist anzumerken, dass das ESTV anstelle des BIPs auch schon die Produktionslücke verwendet hat, was unserer Erfahrung nach aber keinen bedeutenden Unterschied macht (vgl. ESTV 2015).⁴ Relativ dazu wird auf zweifache Weise eine Sophistizierung vollzogen: zum einen wird der Zusammenhang auf der Basis desaggregierter Daten modelliert (kantonale Daten und zusätzliche Berücksichtigung von Einkommensklassen bzw. Steuerstatus), zum anderen werden im Rahmen der Desaggregation auch alternative Indikatoren zu den im aktuellen Schätzmodell verwendeten in Betracht gezogen.

Der Vorteil dieses Untersuchungsansatzes besteht darin, dass durch die schrittweise Sophistizierung herausgeschält werden kann, wieviel Potenzial in jedem Schritt vorhanden ist. Dies nicht zuletzt deshalb, weil die in den einzelnen Schritten verwendeten Methoden und Daten miteinander vergleichbar sind.

⁴ In unseren Regressionen, welche weiter unten erläutert werden, sind die Ergebnisse für das BIP und die Produktionslücke jeweils sehr ähnlich ausgefallen. Es dürfte daher keinen grossen Unterschied machen, welcher der beiden Indikatoren verwendet wird.

2.2 Verwendete Daten

2.2.1 Steuererträge

Grundlagen

Die Grundlage für die Untersuchung des Zusammenhangs zwischen ausgewählten Indikatoren und den Steuererträgen bilden steuerseitig die (online verfügbaren) Steuerstatistiken der Eidgenössischen Steuerverwaltung ESTV.

Für juristische Personen sind in der Steuerstatistik vergleichbare Daten auf nationaler und kantonaler Ebene für den Zeitraum von 1999 bis 2012 erhältlich. Desaggregiert nach dem Steuerstatus (Ordentliche, Holding, Domizil und Gemischte Gesellschaften) liegen diese Daten jedoch nur für den Zeitraum von 2006 bis 2012 vor. In unserer Analyse wurden sowohl die nicht nach dem Steuerstatus aufgeteilten als auch die nach dem Steuerstatus aufgeteilten Daten verwendet.

Die Ergebnisse der direkten Bundessteuer für natürliche Personen nach einjähriger Veranlagung sind in der (online verfügbaren) Steuerstatistik je nach Kanton für unterschiedliche Zeiträume verfügbar, da nicht alle Kantone gleichzeitig von der zweijährigen auf die einjährige Veranlagung umgestellt haben. Für drei Kantone erstreckt sich der Zeitraum von 1999 bis 2012 (Zürich, Basel Stadt, Thurgau), für 20 Kantone von 2001 bis 2012 und für drei Kantone (Tessin, Waadt, Wallis) von 2003 bis 2012. Dies hat entsprechend zur Folge, dass einheitliche Daten für alle Kantone nur für den Zeitraum 2003 bis 2012 verfügbar sind. Die Steuerertragsdaten liegen in der Steuerstatistik für die erwähnten Zeiträume auch desaggregiert nach Nicht-Normalfällen⁵ und Normalfällen vor, wobei die Normalfälle weiter in 21 Einkommensklassen aufgeteilt sind. Aus diesen 21 Einkommensklassen haben wir zu Analyse Zwecken zwei Einkommensklassen gebildet: niedrige Einkommen (unter 200'000 CHF) und hohe Einkommen (200'000 CHF und darüber). Bei der Bildung von Einkommensklassen zu Analyse Zwecken haben wir verschiedene Möglichkeiten in Betracht gezogen. Da eine weitere Unterteilung der Einkommensklassen unter 200'000 CHF ähnliche Resultate lieferte, haben wir in der weiteren Analyse darauf verzichtet, die niedrigen Einkommen weiter zu unterteilen.

Korrektur um Steuerreformen

Um den statistischen Zusammenhang zwischen ökonomischen Variablen und dem Steuerertrag zu überprüfen, ist es wichtig, dass die Entwicklung der Erträge nicht in einzelnen Jahren durch die Auswirkungen von Steuerreformen verzerrt ist, da ansonsten die Wahrscheinlichkeit steigt, dass zufällige Zusammenhänge zwischen Indikatoren und Erträgen identifiziert werden. Gemäss den von der ESTV gelieferten Angaben zu den Auswirkungen der Steuerreformen⁶ betreffen alle grösseren Reformmassnahmen der letzten Jahre die Steuererträge der natürlichen Personen. Daher wurde die Zeitreihe der Steuererträge der natürlichen Personen um den Einfluss der Steuerreformen korrigiert. Hierfür wurde in Jahren mit Steuerreformen der tatsächliche Steuerertrag der natürlichen Personen um die von der ESTV geschätzten Einbus-

⁵ Die Nicht-Normalfälle umfassen: Sonderfälle (Pauschal Besteuerte, Übrige Sonderfälle), Quellensteuer und Kapitalleistungen aus Vorsorge.

⁶ Berücksichtigte Steuerreformen: Ausgleich kalte Progression (2006, 2007, 2011, 2012), Sofortmassnahmen Ehepaar- und Familienbesteuerung (2008), Kapitaleinlageprinzip (2011), Familienbesteuerung (2011).

sen durch die Reform erhöht. Dabei wurde mangels genauerer Informationen unterstellt, dass die Erträge nach Kantonen und Einkommenssteuerklassen von den Reformen in gleichem Masse betroffen sind. In den Jahren ohne Steuerreformen wurden die erhaltenen fiktiven Steuererträge mit den Wachstumsraten der tatsächlichen Steuererträge fortgeschrieben.

Bei der Zeitreihe der Steuererträge der juristischen Personen wurden keine Korrekturen durchgeführt und die tatsächlichen Werte für die statistischen Schätzungen verwendet, da im betrachteten Zeitraum keine Reformen mit wesentlichen Auswirkungen auf die Gewinnsteuer-Entwicklung stattgefunden haben.

2.2.2 Indikatoren aus den BAKBASEL Datenbanken

Für die Auswahl von geeigneten ökonomischen Indikatoren hat BAKBASEL hauptsächlich auf die umfangreiche BAKBASEL Modellwelt zurückgegriffen. Die Schweizer Makro-, Branchen-, und Regionalmodelle enthalten die Komponenten der volkswirtschaftlichen Gesamtrechnung und werden für empirisch fundierte Prognosen verwendet. Daher liegen bei allen verwendeten Indikatoren die Daten nicht nur für historische Zeiträume vor, sondern es sind auch Prognosen bis zum Jahr 2040 verfügbar, welche regelmässig aktualisiert werden. Folgende Variablen wurden als mögliche erklärende Indikatoren in die Untersuchung einbezogen:

- Nominales BIP auf Schweiz- und auf Kantonsebene, geschätzte nominale Outputlücke auf Schweiz- und auf Kantonsebene
- Nominale Bruttowertschöpfung der Gesamtwirtschaft, des Business-Sektors⁷ sowie ausgewählter Branchen und Branchenaggregate⁸, jeweils auf Schweizer und auf Kantonsebene
- Lohnsumme der Gesamtwirtschaft sowie ausgewählter Branchen bzw. Branchenaggregate auf Schweizer und auf Kantonsebene, unbereinigt und bereinigt um Pendlerverflechtungen⁹
- Zahl der Beschäftigten in Vollzeitäquivalenten (unbereinigt und bereinigt um Pendlerverflechtungen), Zahl der Arbeitslosen auf Schweizer und Kantonsebene
- Nominale Schweizer Güter-, Dienstleistungs- und Gesamtexporte sowie nominale Güterexporte der Chemie-/Pharmabranche (A2021) und der Investitionsgüterindustrie (A2430)
- CHF/EUR-Wechselkurs, handelsgewichteter nominaler Franken-Aussenwert
- Schweizer Aktienindex SPI, globaler Aktienindex MSCI World, Renditen zehnjähriger Schweizer Staatsanleihen

⁷ Business Sektor: Aggregat der Branchen 01-82 gemäss NOGA 08 Klassifikation

⁸ Liste der ausgewählten Branchen gemäss NOGA 08 Klassifikation: 05-43, 10, 10-33, 20-21, 24-30, 45-47, 45-98, 64, 64-66, 65, 66, 69-75, 84-88

⁹ Die Lohnsummen und die Zahl der Beschäftigten in Vollzeitäquivalenten auf Kantonsebene wurden mittels einer Pendlermatrix um die Pendlerströme bereinigt, um der Tatsache Rechnung zu tragen, dass für die Einkommensteuer der Wohnort und nicht der Arbeitsort von Bedeutung ist.

- Rohstoff- und Energiepreisindizes
- Proxyvariable für den Gewinn des Business-Sektors¹⁰ auf Schweizer und auf Kantonsebene
- Primäreinkommen auf Ebene Schweiz

Insgesamt haben wir inklusive aller Branchen- und Kantonsvarianten dieser Variablen über 2000 Indikatoren erhalten und in die Untersuchung einbezogen. In dieser Zahl ist noch nicht berücksichtigt, dass wir bei jedem Indikator in den ersten Analyseschritten neben der nicht-verzögerten Version auch ein-, zwei- und dreijährige Lags berücksichtigt haben; in späteren Analyseschritten haben wir neben der nicht-verzögerten Version noch ein einjähriges Lags berücksichtigt, weil sich gezeigt hat, dass höhere Lags keine besseren Ergebnisse bringen.

2.3 Methodologische Vorgehensweise

Unsere Erläuterung der Methoden beginnt im Abschnitt 2.3.1 mit der Einführung des Regressionsmodells, welches der Untersuchung des Zusammenhangs zwischen den makroökonomischen Indikatoren und den Sollerträgen zugrunde liegt. Danach erläutern wir in den Abschnitten 2.3.2 bis 2.3.5 die aufeinander aufbauenden Analyseschritte und die dabei verwendeten Methoden.

2.3.1 Regressionsmodell

Der Grossteil der statistischen Analyse wurde mit der Programmier- und Statistiksoftware R durchgeführt. Dabei zeigte eine deskriptive Analyse der Daten, dass die Zeitreihen der Steuererträge und Indikatoren in der Regel nicht-stationär sind. Um möglichen Problemen vorzubeugen, welche entstehen können, wenn man nicht-stationäre Zeitreihen aufeinander regressiert (spurious regression), wurden für alle vorliegenden Zeitreihen jährliche Wachstumsraten gebildet. Der Nachteil dieser Vorgehensweise ist, dass für die Schätzung ein Jahr weniger zur Verfügung steht.

Das Grundmodell, welches unseren Regressionsanalysen und Backtests der ausgewählten Indikatoren zugrunde liegt, hat sowohl auf nationaler als auf kantonaler Ebene die folgende Form:

$$\Delta y_t = a + bL\Delta x_t + u_t$$

- a: Konstante
- y_t: Sollertrag aus Gewinn- bzw. Einkommenssteuern, bereinigt um Steuerreformen (Zeitreihe)
- x_t: Ökonomische Indikatoren (Zeitreihe)
- b: Koeffizient von x_t (Steigerungsrate)
- u_t: Fehlerterm (Zeitreihe)
- t: Zeit in Jahren. Bei den juristischen Personen: von 2000 bis 2012 falls ohne Desaggregation nach Steuerstatus, von 2007 bis 2012 falls mit Desaggregation

¹⁰ Um eine Zeitreihe mit dem geschätzten Gewinn des Business Sektors zu erhalten, wurde die Lohnsumme von der nominalen Bruttowertschöpfung auf Branchenebene (2-Steller Branchen gemäss NOGA 08) subtrahiert und die erhaltenen positiven Werte über alle Branchen aufaddiert.

tion nach Steuerstatus. Bei den natürlichen Personen: von 2000, 2002 oder 2004 bis 2012, je nachdem, um welche regionale Einheit es sich handelt (vgl. Abschnitt 2.2.1). Die Idee war hier, für jedes Modell die maximale Anzahl der Jahre zu verwenden, um alle verfügbaren Informationen einfließen zu lassen. Zu beachten ist, dass aufgrund der Wachstumsratenbildung gegenüber den in Abschnitt 2.2 erwähnten Zeiträumen ein Jahr verloren geht.

- Δ: Wachstumsrate in Prozent
- L: Zeitliche Verzögerung der Wachstumsrate des Indikators: L0 = keine Verzögerung, L1 = Verzögerung um 1 Jahr. Weitere Verzögerungen wurden anfänglich getestet, aber für die weitere Analyse nicht berücksichtigt, da sie überwiegend einen geringeren Erklärungswert aufwiesen als L0 und L1.

Bei der Untersuchung des Zusammenhangs zwischen ausgewählten Indikatoren und den Steuererträgen von juristischen Personen wurde das Modell auch noch in einer leicht abgewandelten Form geschätzt, worauf an entsprechender Stelle eingegangen wird.

2.3.2 Schätzung der Benchmarks

Der im Abschnitt 2.1 erläuterten Untersuchungsstrategie folgend, haben wir zunächst Benchmarks ermittelt, relativ zu denen beurteilt werden kann, inwiefern Verfeinerungen der Schätzmethode (Desaggregation und alternative Indikatoren) die ex-post Prognosequalität verbessern.

Benchmark B1: CH Sollertrag, CH Indikatoren des gegenwärtigen Schätzmodells, mit Gleichsetzung Wachstumsraten

Bei der ersten Benchmark handelt es sich um die „einfachste“ Modellierung des Zusammenhangs zwischen makroökonomischen Indikatoren und den Sollerträgen. Das heisst: Als Indikatoren werden die im gegenwärtigen Schätzmodell verwendeten Indikatoren verwendet (Primäreinkommen bei der Einkommenssteuer und BIP bei der Gewinnsteuer). Der Zusammenhang wird auf der Basis von aggregierten bzw. nationalen Daten modelliert und es wird vorausgesetzt, dass die Wachstumsraten der Indikatoren den Wachstumsraten der dazugehörigen Sollerträge entsprechen (übersetzt ins Regressionsmodell 2.3.1 heisst das: Konstante $a = 0$, Koeffizient $b = 1$).

Benchmark B2: CH Sollertrag, CH Indikatoren des gegenwärtigen Schätzmodells, mit ökonometrisch geschätzten Koeffizienten

Bei der Benchmark B2 handelt es sich um eine Verfeinerung der Benchmark B1. Das bedeutet: Es werden immer noch die gleichen Indikatoren verwendet wie bei der Benchmark B1 (Primäreinkommen bei der Einkommenssteuer, BIP bei der Gewinnsteuer) und der Zusammenhang wird immer noch auf aggregierter bzw. nationaler Ebene modelliert. Im Unterschied zur Benchmark B1 wird der Zusammenhang zwischen den Indikatoren und den Sollerträgen aber mit dem oben eingeführten Regressionsmodell ökonometrisch geschätzt (das heisst: $a \neq 0$, $b \neq 1$).

Vor dem Hintergrund des Untersuchungsziels in diesem Kapitel, den Nutzen der Desaggregation bei der Modellierung des Zusammenhangs zwischen makroökonomischen Indikatoren und Sollerträgen zu untersuchen, erachten wir die beiden Bench-

marks nicht als gleichwertig: die Benchmark B1 ist einfacher oder „naiver“ als die Benchmark B2. Deshalb halten wir den Vergleich mit der Benchmark B2 für am aussagekräftigsten und verwenden nachfolgend vorrangig diesen Benchmark für die Vergleiche.

2.3.3 Schätzung des Modells für eine Vielzahl von Steuerertrags-Indikatoren-Kombinationen

In einem nächsten Schritt wurde das obige Regressionsmodell für etwa 200'000 Kombinationen von y-Variablen (Steuererträge natürlicher und juristischer Personen auf kantonaler und/oder nationaler Ebene) und x-Variablen (Indikatoren auf kantonaler und/oder nationaler Ebene mit verschiedenen zeitlichen Verzögerungen bzw. Lags) geschätzt. Für jede Kombination wurde zusätzlich zur Schätzung des Regressionsmodells der Korrelationskoeffizient berechnet.

Das Ziel dieses Analyseschritts bestand darin, eine möglichst grosse Datengrundlage zu schaffen (Bestimmtheitsmass, p-Wert des Koeffizienten, Korrelation für jede Steuerertrags-Indikatoren-Kombination), auf welcher in einem nächsten Schritt eine statistisch bzw. ökonomisch begründete Selektion von Indikatoren vollzogen werden konnte.

2.3.4 Auswahl geeigneter Indikatoren-Sets

Als Nächstes wurden aus allen verfügbaren Erklärungsvariablen Sets von geeigneten Indikatoren ausgewählt.

Die Grundidee hinter diesem Analyseschritt und der Verwendung des Begriffs „Indikatoren-Set“ lässt sich folgendermassen erklären: Was uns am Ende interessiert, sind geschätzte Zeitreihen für den Sollertrag von natürlichen bzw. juristischen Personen auf nationaler Ebene, welche man mit den tatsächlichen Werten vergleichen kann, um Erkenntnisse über die ex-post Prognosegüte zu erlangen. Dem Desaggregations-Moment dieser Studie folgend, werden die geschätzten nationalen Zeitreihen als Summen von geschätzten desaggregierten Zeitreihen berechnet: das sind kantonale Zeitreihen (falls nur nach Kantonen desaggregiert wird) oder Zeitreihen mit einer kantonalen Dimension und einer Einkommensklassen- bzw. Steuerstatus-Dimension (falls bei den natürlichen bzw. juristischen Personen noch weiter desaggregiert wird). Die Aufgabe in diesem Analyseschritt bestand darin, für jede dieser Desaggregationsmöglichkeiten ein jeweils geeignetes Set von Indikatoren auszuwählen.

Diesen Punkt kann man auch noch mit Hilfe eines Beispiels erläutern. Betrachten wir die Modellierung des Zusammenhangs zwischen makroökonomischen Indikatoren und den Sollerträgen aus der Einkommenssteuer auf der Basis einer regionalen Desaggregation (ohne zusätzliche Desaggregation nach Einkommensklassen). Dazu muss man 26 Regressionen durchführen, auf deren Basis man 26 Zeitreihen für geschätzte kantonale Sollerträge berechnen kann, deren Summe die geschätzte Zeitreihe für den Schweizer Steuerertrag aus den natürlichen Personen darstellt. Für jede dieser 26 Regressionen muss aus einem grossen Set potenzieller Indikatoren (vgl. Abschnitt 2.2 und 2.3.3) ein geeigneter Indikator (x_t im obigen Regressionsmodell) ausgewählt werden. Diese 26 Indikatoren zusammen bezeichnen wir als Indikatoren-Set.

Aufgrund der grossen Anzahl von möglichen Indikatoren-Sets war es wichtig, methodisch vorzugehen. Dazu haben wir unseren Auswahlprozess sowohl bei den natürlichen als auch bei den juristischen Personen in fünf thematisch aufeinander aufbauende Teilschritte bzw. Schätzmethode gegliedert. Um Missverständnissen vorzubeugen: bei den einzelnen Schätzmethode handelt es sich weniger um verschiedene statistische Verfahren als um verschiedene Arten von Indikatoren-Sets. Bei jeder Schätzmethode wurde einmal nur nach Kantonen desaggregiert, und einmal nach Kantonen und Einkommensklassen bzw. Steuerstatus (jede Schätzmethode resultiert deshalb in mindestens zwei Schätzungen, bei manchen Schätzmethode wurden auch noch weitere Varianten berechnet). Die verschiedenen Schätzmethode werden nun der Reihe nach erläutert.¹¹

Schätzmethode S1: kantonale Sollerträge, CH Indikatoren des gegenwärtigen Schätzmodells

Bei der ersten Schätzmethode wurde der Zusammenhang zwischen den kantonalen Gewinn- bzw. Einkommenssteuererträgen und den im aktuellen Schätzmodell verwendeten nationalen Indikatoren (nominales BIP Schweiz bzw. Primäreinkommen Schweiz) modelliert. In Ergänzung wurde in einer zweiten Variante der Schätzmethode 1 der Zusammenhang zusätzlich nach Einkommensklassen bzw. Steuerstatus differenziert geschätzt. In diesem Schritt sollte somit untersucht werden, ob einzig durch eine Umstellung der Schätzmethode auf einen Bottom-Up-Ansatz eine höhere Prognosequalität erzielt werden kann.

Schätzmethode S2: kantonale Sollerträge, kantonale Indikatoren des gegenwärtigen Schätzmodells

Bei der zweiten Schätzmethode wurde der Zusammenhang zwischen den kantonalen Steuererträgen und den kantonalen Versionen der im aktuellen Modell verwendeten Indikatoren modelliert. Auch hier wurde in einer zweiten Variante der Zusammenhang wieder zusätzlich nach Einkommensklassen bzw. Steuerstatus differenziert.

Als Indikatoren für die kantonalen Sollerträge aus der Gewinnsteuer wurden folglich die kantonalen nominalen BIPs verwendet. Als Indikator für die kantonalen Sollerträge aus der Einkommenssteuer konnten nicht die kantonalen Primäreinkommen verwendet werden, da BAKBASEL die kantonalen Primäreinkommen auf Grundlage von Steuerdaten schätzt. Als Alternative wurden daher die kantonalen Lohnsummen bereinigt um die Pendlerströme verwendet, um der Tatsache Rechnung zu tragen, dass für die Einkommenssteuer der Wohnort und nicht der Arbeitsort von Bedeutung ist.¹²

Schätzmethode S3: Kantonale Sollerträge, Basis-Indikatoren-Sets

In der Schätzmethode S3 wurden verschiedene Basis-Indikatoren-Sets ausgewählt, welche auf Grundlage theoretischer Überlegungen als besonders geeignet für die Modellierung der kantonalen Sollerträge erscheinen. Wir sprechen hier von „Basis-Indikatoren-Sets“, weil wir diese in der Schätzmethode S4 punktuell mit anderen Indikatoren ergänzen.

¹¹ Die tatsächlich verwendeten Indikatoren-Sets wurden den Auftraggebern in einem Excel-File mitgeliefert.

¹² Dies entspricht dem Vorgehen von BAKBASEL bei der Erstellung von Prognosen für die kantonalen Primäreinkommen.

Bei den Gewinnsteuern wurden folgende Varianten von Basis-Indikatoren-Sets in Betracht gezogen:

- Variante 1: Indikatoren-Set bestehend aus dem Gewinn des Business-Sektors auf Kantonsebene.
- Variante 2: Indikatoren-Set bestehend aus folgenden acht Indikatoren: Proxyvariable Gewinn Business-Sektor CH und Kantone, nominales BIP CH und Kantone, nominale Bruttowertschöpfung Gesamtwirtschaft CH und Kantone sowie nominale Bruttowertschöpfung Business-Sektor CH und Kantone; alle ohne zeitliche Verzögerung (= L0). Für jeden Kanton bzw. Steuerstatus wurde derjenige Indikator ausgewählt, welcher das höchste Bestimmtheitsmass aufwies. Dabei wurde nicht berücksichtigt, ob das Vorzeichen der Korrelation den theoretischen Erwartungen entspricht.
- Variante 3: Indikatoren-Set analog zur zweiten Variante, nur wurden zusätzlich auch um ein Jahr verzögerte Variablen (= L1) berücksichtigt.

Bei der Einkommenssteuer wurden zwei Varianten von Basis-Indikatoren-Sets in Betracht gezogen:

- Variante 1: Indikatoren-Set bestehend aus den folgenden sechs Indikatoren: Primäreinkommen CH, Lohnsumme CH, nominales BIP CH und Kantone, Bruttowertschöpfung CH und Kantone; alle ohne zeitliche Verzögerung (= L0). Für jeden Kanton bzw. Einkommensklasse wurde dabei der Indikator gewählt, welcher das höchste Bestimmtheitsmass aufwies. Dabei wurde nicht berücksichtigt, ob das Vorzeichen der Korrelation den theoretischen Erwartungen entspricht.
- Variante 2: Indikatoren-Set analog zur ersten Variante, bei dem die BIP und Bruttowertschöpfungsindikatoren zusätzlich mit einer Verzögerung um ein Jahr berücksichtigt wurden (= L1).

Schätzmethode S4: Kantonale Sollerträge, Basis-Indikatoren-Sets ergänzt mit ausgewählten ökonomischen Indikatoren

Bei der vierten Schätzmethode wurden die Basis-Indikatoren-Sets der Schätzmethode S3 als Ausgangspunkt verwendet und bei einigen Kantonen die Indikatoren der Sets durch ausgewählte andere Indikatoren ausgetauscht.¹³ Voraussetzung war, dass die gewählten alternativen Indikatoren eine ökonomisch nachvollziehbare Interpretation erlauben (z.B. positive Korrelation von Branchenwertschöpfung im Finanzsektor mit Sollerträgen aus der Gewinnsteuer) und zugleich statistisch signifikantere Ergebnisse (= höheres Bestimmtheitsmass) als die Indikatoren des Basis-Indikatoren-Sets lieferten; zusätzlich wurden nur diejenigen Indikatoren in Betracht gezogen, für die auch sinnvolle Prognosen möglich sind.¹⁴

¹³ Diese Indikatoren stammen aus dem Gesamtindikatorset von über 2000 Indikatoren (inkl. branchen- und kantonalen Versionen), zusätzlich mit verschiedenen Lag-Strukturen.

¹⁴ Anders als bei den bisherigen (und den nachfolgenden) Ansätzen spielt hier also auch eine – immer subjektive – Einschätzung des bearbeitenden Ökonomen zur Auswahl der Indikatoren eine Rolle.

Für die Erklärung der Steuererträge der juristischen Personen wurde noch eine Variation dieser Schätzmethode auf der Basis eines abgeänderten (multivariaten) Regressionsmodells durchgeführt. Hierfür wurden beim Indikatoren-Set der Schätzmethode S4 die finanzplatzbezogenen Indikatoren ausgetauscht, aber das Regressionsmodell um einen finanzplatzbezogenen Interaktionsterm ergänzt. Dieser Interaktionsterm besteht aus der Multiplikation von zwei Variablen: Einer Finanzkrisen-Dummyvariable (2001 = 1; 2008 = 1; übrige Jahre = 0) und einem Indikatoren-Set bestehend aus der kantonalen Wertschöpfung des Finanzsektors (Aggregat der NOGA08 Branchen 64 bis 66). Durch diese Variante der Schätzmethode S4 wurde versucht, die Entwicklung der Gewinnsteuererträge in den Finanzkrisenjahren 2001 und 2008 möglichst gut zu erfassen (deshalb der Interaktionsterm) und gleichzeitig die Entwicklung in den Nicht-Finanzkrisenjahren ausschliesslich über nicht-finanzplatzbezogene Indikatoren zu erklären (deshalb die Bereinigung des Indikatoren-Sets um finanzplatzbezogene Indikatoren) (vgl. Abschnitt 2.5.5 für Hintergründe).

Schätzmethode S5: Kantonale Sollerträge, numerisch ausgewählte Indikatoren

Für die letzte Schätzmethode wurde das Indikatoren-Set mechanisch auf Grundlage des je Kanton bzw. je Einkommensklasse/Steuerstatus höchsten Bestimmtheitsmasses ausgewählt ohne auf ökonomisch nachvollziehbare Zusammenhänge und das Vorzeichen der Korrelation zu achten. Allerdings wurde darauf geachtet, dass nur Indikatoren gewählt werden, welche prognostizierbar sind. Dadurch wurden Aktienindizes ohne zeitliche Verzögerung, welche in gewissen Fällen gute Erklärungsvariablen waren, ausgeschlossen.

2.3.5 Backtest der gewählten Indikatoren-Sets

Nachdem die Auswahl der Indikatoren-Sets abgeschlossen war, wurde die geschätzte Entwicklung der Sollerträge auf Bundesebene mit den realisierten Sollerträgen der direkten Bundessteuer verglichen.

Hierfür wurde für alle Schätzmethoden (bzw. Varianten davon) sowie die Benchmarks das gleiche Vorgehen angewendet. Dieses wird der Einfachheit wegen nur für die Desaggregation nach Kantonen erläutert, für die Desaggregation nach Kantonen und Einkommensklassen bzw. Steuerstatus verläuft alles analog:

Zunächst wurden für jeden Kanton auf der Basis des Regressionsmodells und der ausgewählten Indikatoren-Sets Zeitreihen für die Wachstumsrate des Steuerertrags geschätzt; dies entspricht den „fitted values“ aus dem Regressionsmodell:

$$\Delta y_t^f = a + bL\Delta x_t$$

wobei f für „fitted“ bzw. „forecasted“ steht und sich die restliche Terminologie am Modell in Abschnitt 2.3.1 orientiert.

Auf der Basis dieser geschätzten Wachstumsraten wurden anschliessend für jeden Kanton geschätzte Zeitreihen für die Levels des Steuerertrags berechnet. Bei dieser Berechnung wurden die Wachstumsraten jeweils auf dem Level des Vorjahres aufgesetzt, das heisst:

$$y_t^f = y_{t-1} * (1 + \Delta y_t^f)$$

Als Nächstes wurden diese Zeitreihen über die Kantone aggregiert um eine Zeitreihe für den geschätzten Schweizer Steuerertrag zu erhalten. Das heisst:

$$Y_t^f = \sum y_t^f$$

wobei aggregierte bzw. nationale Grössen mit Grossbuchstaben denotiert werden und hier über Kantone summiert wird. Dieser Schritt wurde bei den Benchmarks ausgelassen, weil es sich dabei bereits um Zeitreihen auf der Ebene Schweiz handelt.

Auf der Basis dieser geschätzten Niveauzeitreihen für den Steuerertrag auf der Ebene Schweiz lässt sich schliesslich die Performance oder Prognosequalität der verschiedenen Schätzmethoden und der Benchmarks vergleichen. Dieser Performancevergleich kann neben graphischen Mitteln auf der Basis verschiedener Masse vorgenommen werden. Konkret verwenden wir die Folgenden:

- Durchschnitt der absoluten Abweichungen pro Jahr =

$$\frac{\sum |Y_t - Y_t^f|}{\text{Anzahl Vergleichsjahre}}$$

wobei Y_t für den tatsächlichen Sollertrag der Schweiz in t steht, $|\cdot|$ für die Betragsfunktion und hier über die Vergleichsjahre summiert wird.

- Durchschnitt der quadrierten Abweichungen pro Jahr =

$$\frac{\sum (Y_t - Y_t^f)^2}{\text{Anzahl Vergleichsjahre}}$$

wobei hier wieder über die Vergleichsjahre summiert wird. Die Idee bei diesem Mass ist, dass im Gegensatz zum Mass „Durchschnitt der absoluten Abweichungen pro Jahr“ grössere Abweichungen stärker „bestraft“ werden als kleinere Abweichungen.

- Bestimmtheitsmass der Regression der Wachstumsrate des tatsächlichen Sollertrags auf die Wachstumsrate des geschätzten Sollertrags:

$$\Delta Y_t = a + b\Delta Y_t^f + u_t$$

wobei die Notation weiter oben eingeführt wurde.

2.4 Ergebnisse für die Einkommenssteuer

2.4.1 Überblick

Das Ziel unserer Analyse des Zusammenhangs zwischen makroökonomischen Indikatoren und dem Sollertrag aus der Einkommenssteuer ist es, herauszufinden, ob mithilfe von desaggregierten Daten und Schätzungen eine Verbesserung der ex-post Prognosequalität erzielt werden kann. Im Folgenden werden die Abweichungen der verschiedenen zur Diskussion stehenden Sollertragsreihen von den tatsächlichen Sollertragsreihen einander gegenübergestellt. Die Ergebnisse sind in der nachfolgenden

den Tabelle zusammengefasst und werden in den nächsten Abschnitten der Reihe nach erläutert.

Tab. 2-1 Vergleich der ex-post Prognosequalität von verschiedenen Benchmarks und Schätzmethoden für die Sollerträge der Einkommenssteuer

Benchmarks (B) bzw. Schätzmethode (S)	Durchschnitt der absoluten Abweichungen pro Jahr		Durchschnitt der quadrierten Abweichungen pro Jahr		Bestimmtheitsmass
	Abweichung in Franken	% Veränderung ggü. Abweichung B2	Abweichung in Franken	% Veränderung ggü. Abweichung B2	
B1 CH Sollertrag, CH Primäreinkommen, mit Gleichsetzung Wachstumsraten	3.87E+08	57%	2.19E+17	45%	0.00
B2 CH Sollertrag, CH Primäreinkommen, mit ökonomisch geschätzten Koeffizienten	2.46E+08	0%	1.51E+17	0%	0.09
S1 Kant. Sollertrag, CH Primäreinkommen	2.44E+08	-1%	1.63E+17	8%	0.05
Berücksichtigung der Einkommensklassen	2.42E+08	-2%	1.96E+17	30%	0.06
S2 Kant. Sollertrag, Kant. Lohnsumme (pendlerbereinigt)	3.39E+08	38%	1.78E+17	18%	0.11
Berücksichtigung der Einkommensklassen	3.22E+08	31%	2.14E+17	42%	0.10
S3 Kant. Sollertrag, Basis-Indikatoren-Set	2.02E+08	-18%	0.85E+17	-44%	0.45
Berücksichtigung der Einkommensklassen	1.90E+08	-23%	1.00E+17	-34%	0.53
S4a Kant. Sollertrag, Basis-Indikatoren-Set ergänzt um ausgewählte Indikatoren	1.93E+08	-22%	0.81E+17	-46%	0.52
Berücksichtigung der Einkommensklassen	2.27E+08	-8%	1.10E+17	-27%	0.44
S4b Kant. Sollertrag, Basis-Indikatoren-Set stärker ergänzt um ausgewählte Indikatoren	1.75E+08	-29%	0.58E+17	-62%	0.69
Berücksichtigung der Einkommensklassen	1.67E+08	-32%	0.75E+17	-50%	0.59
S5 Kant. Sollertrag, Numerisch ausgewählte Indikatoren	1.34E+08	-46%	0.33E+17	-78%	0.80
Berücksichtigung der Einkommensklassen	1.74E+08	-29%	0.52E+17	-66%	0.82

Bemerkungen: Bei B2 und S1 bis S5 wurden die Koeffizienten der zugrundeliegenden Regressionen ökonomisch geschätzt. Der Vergleichszeitraum umfasst bei allen Benchmarks und Schätzungen (inkl. der Berücksichtigung von Einkommensklassen) die Jahre 2004 bis 2012; deshalb sind alle Werte in der Tabelle direkt miteinander vergleichbar. Die Spalten mit der Bezeichnung „Abweichung in Franken“ geben die durchschnittliche jährliche (absolute bzw. quadrierte) Abweichung der geschätzten Werte von den tatsächlichen Werten in Franken an (vgl. Text für die Formel); die dabei verwendete E-Notation kann so erläutert werden: $1.00E+09 = 1.00 * 10^9 = 1 \text{ Mia}$. Die Spalten mit der Bezeichnung „% Veränderung ggü. der Abweichung B2“ geben an, wie sich der entsprechende Prognosefehler relativ zum Prognosefehler der Benchmark B2 verändert; das heisst, um wieviel Prozent die durchschnittliche jährliche (absolute bzw. quadrierte) Abweichung in Franken gegenüber der Abweichung bei Benchmark B2 reduziert/verbessert (negative Prozentzahlen) bzw. gesteigert/verschlechtert wird (positive Prozentzahlen).

Quelle: BAKBASEL

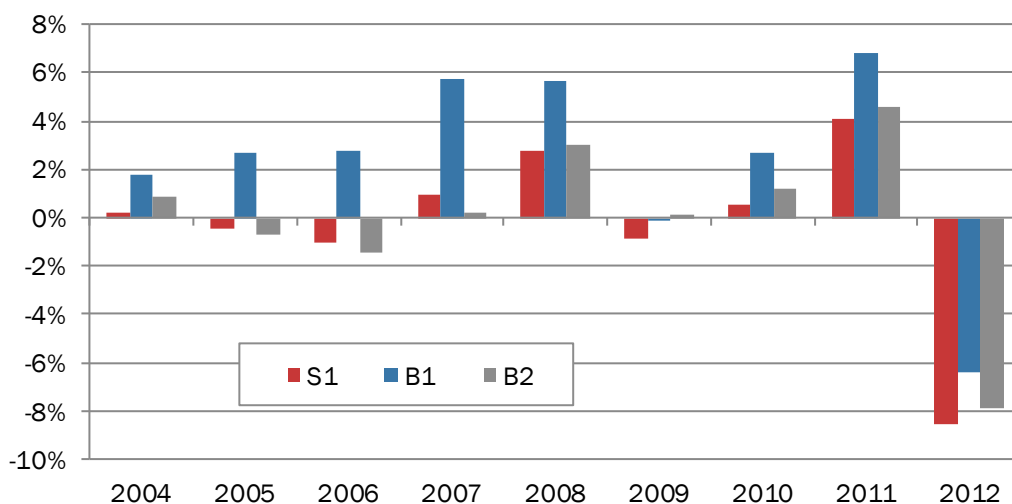
2.4.1 Schätzmethode S1

In der ersten Schätzung wurde untersucht, ob die Prognosequalität im Backtest allein schon dadurch verbessert werden kann, dass der im aktuellen Schätzmodell verwendete Indikator (Primäreinkommen Schweiz) als erklärende Variable für die Schätzung

der kantonalen Sollerträge der Einkommenssteuer gewählt wird. Dies ist jedoch nur begrenzt der Fall: die Schätzmethode S1 schneidet etwas besser ab als die Benchmark 1, aber schlechter als die relevantere Benchmark 2.

Die folgende Abbildung stellt die jährlichen Abweichungen für die Schätzmethode S1 und die beiden Benchmarks in Prozent des Schätzwertes graphisch dar. Auffällig ist, dass alle drei Prognosen (Schätzmethode S1 und beide Benchmarks) insbesondere in 2011 und 2012 relativ weit daneben lagen; in etwas geringerem Ausmass gilt dies auch für 2008.

Abb. 2-1 S1: Abweichung des geschätzten Steuerertrags pro Jahr



Anmerkung: Abweichung in Prozent des Schätzwertes; ein positiver Wert entspricht einer Unterschätzung des tatsächlichen Steuerertrags, ein negativer Wert einer Überschätzung.
Quelle: BAKBASEL

Die Berücksichtigung der zusätzlichen Informationen über die Einkommensklassen (Unterscheidung zwischen Nicht-Normalfällen und Normalfällen, mit einer Unterteilung der Normalfälle in eine tiefe (unter 200'000 CHF) und hohe Einkommensklasse (200'000 CHF und darüber), vgl. Abschnitt 2.2.1 für Details) bringt keine wirkliche Verbesserung der Prognoseperformance. Die jährlichen Abweichungen bei der Berücksichtigung der Einkommensklassen fallen sowohl hier als auch bei den anderen Schätzmethoden sehr ähnlich aus wie ohne deren Berücksichtigung, weshalb auf eine graphische Darstellung dieser Abweichungen verzichtet wird.

2.4.2 Schätzmethode S2

Bei der Schätzung S2 ging es darum, die kantonale Version der Indikatoren im aktuellen Schätzmodell für die Erklärung der Einkommenssteuer zu verwenden. Da die kantonalen Primäreinkommen aus methodologischen Gründen nicht verwendet werden konnten, wurde auf die kantonalen pendlerbereinigten Lohnsummen ausgewichen (vgl. Abschnitt 2.3.4 für Details). Dieser Lohnsummen-Indikator hat aus einer theoretischen Perspektive das höchste Erklärungspotenzial für die Entwicklung der Steuererträge der natürlichen Personen. Aber nicht aus der statistischen Perspektive: wie in der obigen Übersichtstabelle ersichtlich ist, schneidet die Schätzmethode nicht

besser, sondern eher schlechter ab als die Benchmarks. Auch die weitere Desaggregation nach Einkommensklassen bringt keine Verbesserung.

In der folgenden Abbildung sind die jährlichen Abweichungen für die Schätzmethode S2 und die beiden Benchmarks in Prozent des Schätzwertes abgebildet.

Abb. 2-2 S2: Abweichung des geschätzten Steuerertrags pro Jahr



Anmerkung: Abweichung in Prozent des Schätzwertes; ein positiver Wert entspricht einer Unterschätzung des tatsächlichen Steuerertrags, ein negativer Wert einer Überschätzung
Quelle: BAKBASEL

2.4.3 Schätzmethode S3

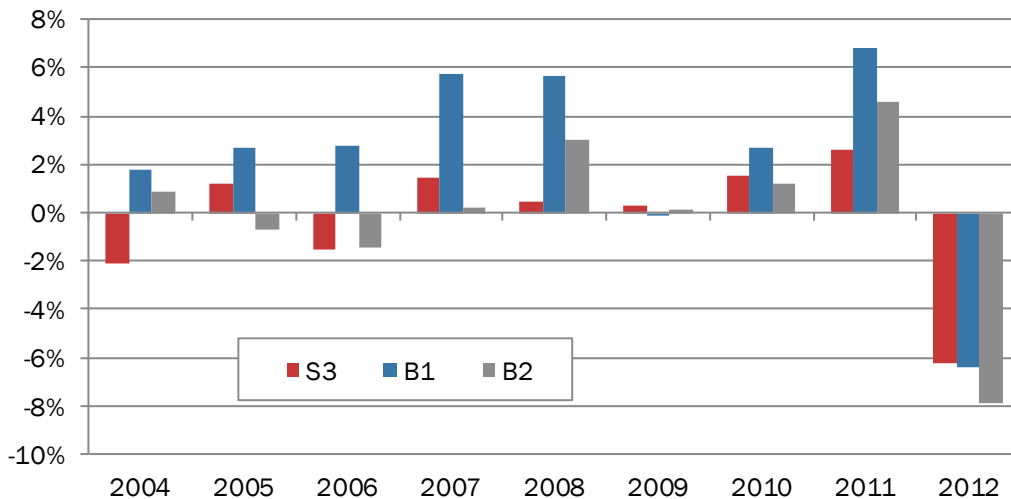
Bei der Schätzmethode S3 wurde getestet, ob mithilfe von kantonalen Alternativ-Indikatoren die Prognosequalität erhöht werden kann. Dabei ging es hier auch darum, Basis-Indikatoren-Sets zu selektionieren, welche bei der Schätzmethode S4 punktuell um ausgewählte weitere Indikatoren ergänzt werden können. Insgesamt haben wir bei der Einkommenssteuer zwei verschiedene Varianten von Basis-Indikatoren-Sets getestet: Die erste Variante des Indikatoren-Sets ist zusammengesetzt aus folgender Gruppe von Indikatoren ohne zeitliche Verzögerung: Primäreinkommen CH, Lohnsumme CH, Nominales BIP CH und Kantone, Bruttowertschöpfung CH und Kantone. Bei der zweiten Variante des Indikatoren-Sets haben wir noch einjährig verzögerte Indikatoren (BIP und Bruttowertschöpfungsindikatoren) ins Set aufgenommen (für Details und die Selektion der Indikatoren, vgl. Abschnitt 2.3.4). Die Berücksichtigung von zeitlich verzögerten Versionen des BIPs und der Bruttowertschöpfung kann Sinn machen, weil diese Variablen unter anderem über die Löhne auf Einkommenssteuer wirken und die Löhne je nach Kontext verzögert auf Veränderungen dieser Variablen reagieren.

Beide Basis-Indikatoren-Sets schneiden gemäss unseren drei Massen zur Beurteilung der Prognosegüte deutlich besser ab als die Benchmarks. Das Basis-Indikatoren-Set, welches unter anderem auch um ein Jahr verzögerte Indikatoren beinhaltet (Variante 1), schneidet dabei etwas besser ab als jenes Set, welches keine verzögerten Indikatoren enthält (Variante 2). Im Folgenden wird deshalb nur noch die

zweite Variante berücksichtigt. Bei dieser zweiten Variante bringt eine weitere Desaggregation nach Einkommensklassen keine klare Verbesserung gegenüber der nicht-desaggregierten Version.

Die folgende Abbildung stellt die jährlichen Abweichungen für die Schätzmethode S3 sowie für die die beiden Benchmarks in Prozent des Schätzwertes graphisch dar.

Abb. 2-3 S3: Abweichung des geschätzten Steuerertrags pro Jahr



Anmerkung: Abweichung in Prozent des Schätzwertes; ein positiver Wert entspricht einer Unterschätzung des tatsächlichen Steuerertrags, ein negativer Wert einer Überschätzung
Quelle: BAKBASEL

2.4.4 Schätzmethode S4a und S4b

Bei der Schätzmethode S4 wurden die Erklärungsvariablen aus der Schätzung S3 in bestimmten Kantonen durch ausgewählte Indikatoren ersetzt. Voraussetzung war, dass die gewählten alternativen Indikatoren eine ökonomisch nachvollziehbare Interpretation erlauben (z.B. positive Korrelation zwischen der pendlerbereinigten Lohnsumme der übrigen Finanzdienstleistungen im Kanton Zug mit den Einkommenssteuererträgen im Kanton Zug), statistisch signifikantere Ergebnisse (= höheres Bestimmtheitsmass) als die Indikatoren der Basis-Indikatoren-Sets liefern sowie prinzipiell prognostizierbar sind.

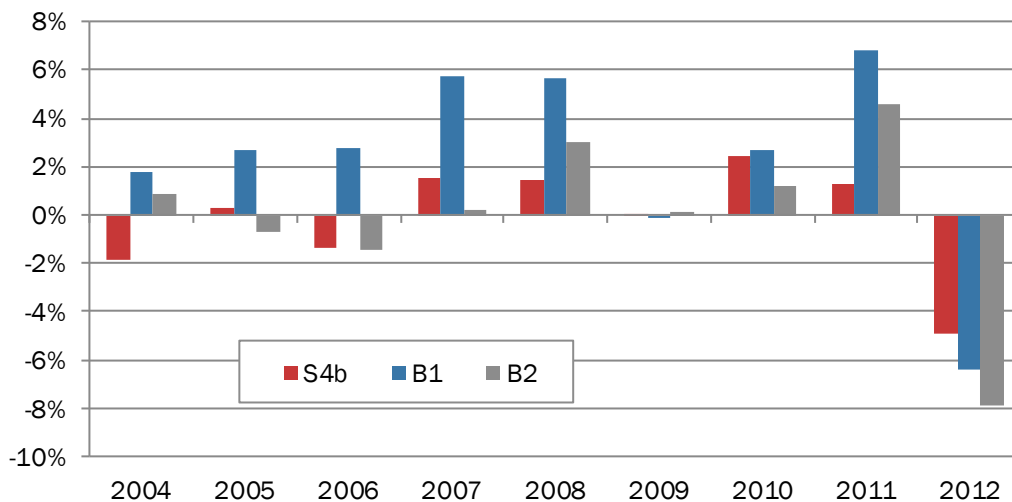
Bei den natürlichen Personen wurden zwei verschiedene Versionen geschätzt. Bei der Schätzmethode S4a wurde das Basis-Indikatoren-Set (Variante 2 von S3, d.h. jene Variante, welche unter anderem verzögerte Werte enthält) nur sehr punktuell ergänzt. Bei der Schätzmethode S4b wurde dasselbe Basis-Indikatoren-Set stärker ergänzt. Der Unterschied zwischen den beiden Versionen S4a und S4b ist, dass die erste Version ökonomisch gesicherter ist, während die zweite Version einen hypothetischen Charakter hat.

Beide Versionen dieser Schätzmethode stellen eine deutliche Verbesserung der Prognosequalität gegenüber den beiden Benchmarks dar. Beide Schätzmethoden

schneiden auch besser ab als die Schätzmethode S3, wobei die Verbesserung bei S4b deutlicher ausgeprägt ist als bei S4a.

Auch hier haben wir wieder eine weitere Desaggregation nach Einkommensklassen in Betracht gezogen. Bei beiden Versionen der Schätzmethode S4 wurden für die Nicht-Normalfälle das unveränderte Basis-Indikatoren-Set (Variante 2 von S3) übernommen. Der Grund dafür ist, dass wir keine zusätzlichen Variablen finden konnten, welche ein höheres Bestimmtheitsmass haben und eine ökonomisch vernünftige Interpretation zulassen. Bei beiden Versionen der Schätzmethode S4 wurde bei den tiefen Einkommen (unter 200'000 CHF) je ein neues Basis-Indikatoren-Set (gegenüber der Variante 2 von S3) hinzugezogen, weil es sich gezeigt hat, dass man bei dieser Einkommensklasse mit dem bestehenden Basis-Indikatoren-Set nicht weiter kommt. Im Fall von S4a wurde als Basis-Indikatoren-Set die um ein Jahr verzögerte pendlerbereinigte Beschäftigung im Dienstleistungssektor gewählt. Im Fall S4b wurde die um ein Jahr verzögerte Lohnsumme im Finanzsektor gewählt. Beide Indikatoren-Sets waren in praktisch allen Kantonen statistisch sehr signifikant. Während das erste Basis-Indikatoren-Set (verzögerte pendlerbereinigte Beschäftigung im Dienstleistungssektor) ökonomisch intuitiv ist, ist das zweite Set eher hypothetischer Natur: es könnte sein, dass die variablen Lohnbestandteile, welche im Finanzsektor auch bei „niedrigen Einkommen“ (unter 200'000 CHF) bedeutsam sind, für die Erklärung der Varianz der Steuererträge in dieser Einkommensklasse den Ausschlag geben. Die einjährige Verzögerung kann möglicherweise damit erklärt werden, dass Unternehmen die variablen Lohnbestandteile in dem Jahr ausweisen müssen, auf dessen Basis sie berechnet werden; und dass die Arbeitnehmer sie erst in dem Jahr versteuern müssen, in welchem sie ausbezahlt werden, was typischerweise erst im darauffolgenden Jahr ist. Für die Klasse hoher Einkommen (200'000 CHF und darüber) wurden bei beiden Schätzmethoden das gleiche Indikatoren-Set verwendet (entspricht Basis-Indikatoren-Set von S3 mit punktuell ausgetauschten Variablen).

Abb. 2-4 S4b: Abweichung des geschätzten Steuerertrags



Anmerkung: Abweichung in Prozent des Schätzwertes; ein positiver Wert entspricht einer Unterschätzung des tatsächlichen Steuerertrags, ein negativer Wert einer Überschätzung
Quelle: BAKBASEL

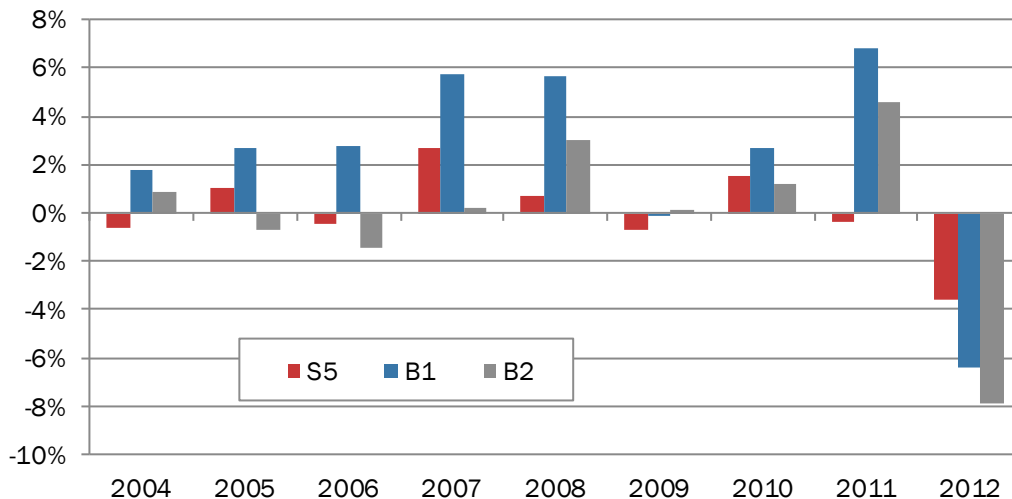
Auch wenn wir bei der nach Einkommensklassen desaggregierten Varianten der Schätzmethoden S4a und S4b an die Grenze dessen gegangen sind, was aus einer ökonomischen Perspektive noch gerechtfertigt werden kann, konnte mit einer zusätzlichen Desaggregation nach Einkommensklassen keine weitere Verbesserung gegenüber den nicht-desaggregierten Varianten erreicht werden.

Die vorangegangene Abbildung 2-4 stellt die jährlichen Abweichungen für die Schätzmethode S4b sowie für die beiden Benchmarks in Prozent des Schätzwertes graphisch dar. Was hier auffällt ist, dass die Abweichungen für das Jahr 2011 und 2012 das erste Mal deutlich kleiner wurden als bei den Benchmarks. Die Ergebnisse für die Schätzmethode S4a sind hier relativ ähnlich und werden deshalb nicht abgebildet.

2.4.1 Schätzmethode S5

Als Letztes wurde bei der Schätzmethode S5 das Indikatoren-Set mechanisch auf Basis des höchsten Bestimmtheitsmasses ausgewählt, ohne auf ökonomisch nachvollziehbare Zusammenhänge und das Vorzeichen der Korrelation zu achten. Wie zu erwarten, wurde mit dieser Schätzmethode die beste ex-post Prognosequalität erreicht. Auch hier macht es aber keinen bedeutsamen Unterschied, ob man eine weitere Desaggregation nach Einkommensklassen vornimmt oder nicht. Die folgende Abbildung stellt die jährlichen Abweichungen für die Schätzmethode S5 sowie für die die beiden Benchmarks in Prozent des Schätzwertes graphisch dar. Auffällig ist, dass gegenüber der Schätzmethode S4 nochmals eine weitere Verbesserung in der Prognosequalität für die Jahre 2011 und 2012 erreicht werden konnte.

Abb. 2-5 S5: Abweichung des geschätzten Steuerertrags



Anmerkung: Abweichung in Prozent des Schätzwertes; ein positiver Wert entspricht einer Unterschätzung des tatsächlichen Steuerertrags, ein negativer Wert einer Überschätzung
Quelle: BAKBASEL

2.5 Ergebnisse für die Gewinnsteuer

2.5.1 Überblick

Das Ziel der Analyse des Zusammenhangs zwischen makroökonomischen Indikatoren und dem Sollertrag aus der Gewinnsteuer besteht darin, herauszufinden, ob mithilfe von desaggregierten Daten und Schätzungen eine Verbesserung der ex-post Prognosequalität erzielt werden kann. In diesem Abschnitt werden die Abweichungen der verschiedenen zur Diskussion stehenden Sollertragsreihen von den tatsächlichen Sollertragsreihen miteinander verglichen. Die Ergebnisse sind in der nachfolgenden Tabelle zusammengefasst und werden im Rest des Abschnitts der Reihe nach erläutert.

Zunächst noch eine Vorbemerkung zur Vergleichbarkeit der Ergebnisse für die Schätzungen ohne Berücksichtigung des Steuerstatus und jenen mit zusätzlicher Berücksichtigung des Steuerstatus. Der Vergleichszeitraum umfasst bei den Schätzungen (S1 bis S5) ohne zusätzlicher Desaggregation nach Steuerstatus die Jahre 2000 bis 2012 (in der Tabelle sind dies die Werte ohne runden Klammern), bei den Schätzungen (S1 bis S5) mit einer zusätzlichen Desaggregation nach Steuerstatus hingegen nur die Jahre 2007-2012 (Werte in runden Klammern). Die Vergleichsmasse für die Benchmarks wurden für beide Vergleichszeiträume berechnet. Diese unterschiedlichen Vergleichszeiträume reduzieren die Vergleichbarkeit der beiden Gruppen von Schätzungen (bzw. die Vergleichbarkeit der Werte ohne runden Klammern mit den Werten in runden Klammern) aus drei Gründen: Erstens waren die Abweichungen in den Jahren 2001, 2005 und 2008 tendenziell gross. Zwei dieser drei Jahre sind aber im kürzeren Beobachtungszeitraum 2007 bis 2012 nicht enthalten. Zweitens steigt das Bestimmtheitsmass mit einer Verkürzung der Zeitreihen an, was das Bestimmtheitsmass bei der Schätzung ohne Berücksichtigung des Steuerstatus praktisch unvergleichbar macht mit dem Bestimmtheitsmass bei der Berücksichtigung des Steuerstatus. Drittens sind Ergebnisse, die auf einem kürzeren Vergleichszeitraum beruhen, im Allgemeinen anfälliger für Zufälle und deshalb weniger belastbar als Vergleiche, die auf längeren Zeiträumen beruhen.

Tab. 2-2 Vergleich der ex-post Prognosequalität von verschiedenen Benchmarks und Schätzmethode für die Sollerträge der Gewinnsteuer

Benchmarks (B) bzw. Schätzmethode (S)	Durchschnitt der absoluten Abweichungen pro Jahr		Durchschnitt der quadrierten Abweichungen pro Jahr		Bestimmtheitsmass
	Abweichung in Franken	% Veränderung ggü. Abweichung B2	Abweichung in Franken	% Veränderung ggü. Abweichung B2	
B1 CH Sollertrag CH BIP, mit Gleichsetzung Wachstumsraten	7.12E+08 (6.41E+08)	7% (-4%)	7.18E+17 (6.20E+17)	-6% (-29%)	0.00 (0.13)
B2 CH Sollertrag, CH BIP, mit ökonomisch geschätzten Koeffizienten	6.66E+08 (6.65E+08)	0% (0%)	7.65E+17 (8.77E+17)	0% (0%)	0.02 (0.10)
S1 Kant. Sollertrag, CH BIP	7.08E+08	6%	9.04E+17	18%	0.01
Berücksichtigung des Steuerstatus	(6.67E+08)	(0%)	(7.65E+17)	(-13%)	(0.63)
S2 Kant. Sollertrag, Kant. BIP	6.78E+08	2%	8.38E+17	10%	0.01
Berücksichtigung des Steuerstatus	(6.31E+08)	(-5%)	(6.89E+17)	(-21%)	(0.56)
S3 Kant. Sollertrag, Basis-Indikatoren-Set	6.78E+08	2%	7.65E+17	0%	0.01
Berücksichtigung des Steuerstatus	(6.28E+08)	(-6%)	(6.71E+17)	(-23%)	(0.59)
S4a Kant. Sollertrag, Basis-Indikatoren-Set ergänzt um ausgewählte Indikatoren	6.94E+08	4%	6.77E+17	-12%	0.07
Berücksichtigung des Steuerstatus	(6.18E+08)	(-7%)	(6.60E+17)	(-25%)	(0.09)
S4b Kant. Sollertrag, Basis-Indikatoren-Set ergänzt um ausgewählte Indikatoren und mit Finanzkrisenterm ¹⁵	5.11E+08	(-23%)	3.38E+17	-56%	0.22
Berücksichtigung des Steuerstatus	(3.26E+08)	(-51%)	(1.40E+17)	(-84%)	(0.41)
S5 Kant. Sollertrag, Numerisch ausgewählte Indikatoren	3.85E+08	-42%	2.53E+17	-67%	0.29
Berücksichtigung des Steuerstatus	(1.06E+08)	(-84%)	(1.70E+17)	(-98%)	(0.89)

Bemerkungen: Bei B2 und S1 bis S5 wurden die Koeffizienten der zugrundeliegenden Regressionen ökonomisch geschätzt. Der Vergleichszeitraum umfasst bei den Benchmarks und den Schätzungen ohne zusätzlicher Desaggregation nach Steuerstatus die Jahre 2000 bis 2012 (Werte ohne runde Klammern), bei den Schätzungen mit einer zusätzlichen Desaggregation nach Steuerstatus hingegen nur die Jahre 2007-2012 (Werte in runden Klammern); die Ergebnisse zwischen den beiden Gruppen sind deshalb nicht ohne Weiteres vergleichbar, weil in 2001 und 2005 die Abweichungen tendenziell gross waren (betrifft alle drei Vergleichsmasse), das Bestimmtheitsmass mit einer Verkürzung der Zeitreihen ansteigt (betrifft nur das Bestimmtheitsmass) und auf kürzeren Zeiträumen basierende Vergleiche im Allgemeinen weniger belastbar sind als auf längeren Zeiträumen beruhende Vergleiche. Die Spalten mit der Bezeichnung „Abweichung in Franken“ geben die durchschnittliche jährliche (absolute bzw. quadrierte) Abweichung der geschätzten Werte von den tatsächlichen Werten in Franken an (vgl. Text für die Formel); die dabei verwendete E-Notation kann so erläutert werden: $1.00E+09 = 1.00 \cdot 10^9 = 1 \text{ Mia}$. Die Spalten mit der Bezeichnung „% Veränderung ggü. der Abweichung B2“ geben an, wie sich der entsprechende Prognosefehler relativ zum Prognosefehler der Benchmark B2 verändert; das heisst, um wieviel Prozent die durchschnittliche jährliche (absolute bzw. quadrierte) Abweichung in Franken gegenüber der Abweichung bei Benchmark B2 reduziert/verbessert (negative Prozentzahlen) bzw. gesteigert/verschlechtert wird (positive Prozentzahlen); dabei werden die Abweichungen der Schätzmethode im Vergleichszeitraum 2000 bis 2012 (Werte ohne runde Klammern) mit den Abweichungen der Benchmarks im Vergleichszeitraum 2002 bis 2012 (Werte ohne runde Klammern) verglichen und analog dazu die Abweichungen der Schätzmethode im Vergleichszeitraum 2007 bis 2012 (Werte in runden Klammern) mit den Abweichungen der Benchmarks im Vergleichszeitraum 2007 bis 2012 (Werte in runden Klammern). Quelle: BAKBASEL

2.5.2 Schätzmethode S1

Bei der ersten Schätzmethode wurde untersucht, ob die Prognosequalität im Backtest allein schon dadurch verbessert werden kann, dass der im aktuellen Modell verwendete Indikator (nominales BIP Schweiz) als erklärende Variable für die Schätzung der kantonalen Gewinnsteuererträge gewählt wird. Dies war jedoch nicht der Fall: der Durchschnitt der absoluten Abweichungen pro Jahr lag bei der Schätzmethode S1 zwar leicht unter der Benchmark 1, aber deutlich über der wichtigeren Benchmark 2. Betrachtet man den durchschnittlichen quadrierten Abweichungswert, dann schnitt die Schätzung S1 sogar deutlich schlechter ab als beide Benchmarks. Die verhaltene ex-post Prognosequalität der Schätzmethode S1 lässt sich dadurch erklären, dass die Wachstumsrate des nominalen BIP der Schweiz in vielen Kantonen nur einen geringen Zusammenhang mit den kantonalen Gewinnsteuererträgen aufweist. In einzelnen Kantonen war sogar eine negative Korrelation zwischen dem Schweizer BIP und den kantonalen Steuererträgen zu beobachten.

Erst die zusätzliche Desaggregation nach dem Steuerstatus bringt gemessen an zwei der drei Vergleichsmasse eine Verbesserung gegenüber der Benchmark B2. Man muss sich aber bewusst sein, dass dieser Vergleich auf einem kürzeren Zeitraum als der Vergleich ohne Desaggregation nach dem Steuerstatus beruht und deshalb nicht gleich belastbar ist (vgl. Abschnitt 2.5.1).

2.5.3 Schätzmethode S2

Bei der Schätzmethode S2 wurde das kantonale nominale BIP als Erklärungsfaktor für die Gewinnsteuererträge verwendet. Gegenüber der Methode S1 konnte dadurch die ex-post Prognosequalität leicht verbessert werden. Allerdings lag der Durchschnittswert der jährlichen absoluten und quadrierten Abweichungen immer noch über den Resultaten der Benchmark 2.

Abb. 2-6 S2: Abweichung des geschätzten Steuerertrags pro Jahr



Anmerkung: Abweichung in Prozent des Schätzwertes; ein positiver Wert entspricht einer Unterschätzung des tatsächlichen Steuerertrags, ein negativer Wert einer Überschätzung
Quelle: BAKBASEL

Durch die Berücksichtigung der zusätzlichen Informationen über den Steuerstatus kann die Prognosequalität erneut verbessert werden, wodurch sowohl die absoluten als auch die quadrierten Abweichungen pro Jahr unterhalb den Ergebnissen der beiden Benchmarks liegen. Aber auch hier gilt es darauf hinzuweisen, dass dieses Ergebnis auf einem kürzeren Vergleichszeitraum beruht und deshalb nicht gleich belastbar ist wie der Vergleich ohne Berücksichtigung des Steuerstatus.

Abb. 2-7 S2 mit Berücksichtigung des Steuerstatus: Abweichung des geschätzten Steuerertrags pro Jahr



Anmerkung: Abweichung in Prozent des Schätzwertes; ein positiver Wert entspricht einer Unterschätzung des tatsächlichen Steuerertrags, ein negativer Wert einer Überschätzung; Quelle: BAKBASEL

2.5.4 Schätzmethode S3

Bei der Schätzmethode S3 wurde getestet, ob mit Hilfe von kantonalen alternativen Indikatoren die ex-post Prognosequalität erhöht werden kann. Hierfür wurden drei Varianten von Basis-Indikatoren-Sets getestet (vgl. Schätzmethode S3 in Abschnitt 2.3.4), deren Ergebnisse allesamt relativ ähnlich ausfielen. In der obigen Tabelle sind die Ergebnisse für die Variante 1 abgebildet, d.h. für ein Basis-Indikatoren-Set bestehend aus dem Gewinn des Business-Sektors auf Kantonsebene. Insgesamt konnte jedoch mit keiner der getesteten Varianten eine bessere Prognosequalität als bei den beiden Benchmarks erzielt werden.

Einzig bei der Berücksichtigung des Steuerstatus wurden leicht bessere Schätzergebnisse erzielt, wobei man sich bewusst sein muss, dass der Vergleichszeitraum kürzer ist und die Ergebnisse deshalb nicht gleich belastbar sind (vgl. Abschnitt 2.5.1).

2.5.5 Schätzmethoden S4a und S4b

Bei der Schätzmethode S4a wurden die Erklärungsvariablen aus der Schätzmethode S3 (genauer: die Variante 1 des Basis-Indikatoren-Sets bestehend aus dem Gewinn des Business-Sektors auf Kantonsebene) in bestimmten Kantonen auf Grundlage

von empirischen Resultaten (Bestimmtheitsmass) und theoretischen Überlegungen durch andere Indikatoren ersetzt. Beispielsweise wurde in einigen Kantonen der handelsgewichtete Wechselkursindex als erklärende Variable verwendet, in einigen anderen Kantonen wurde die Bruttowertschöpfung des sekundären Sektors oder des Finanzsektors als Indikator gewählt. Einige Indikatoren wie der Schweizer Aktienmarkt-Index SPI lieferten zwar ein recht hohes Bestimmtheitsmass in den Regressionen mit den kantonalen Sollerträgen als abhängige Variablen, wurden jedoch aufgrund der unzureichenden Prognostizierbarkeit nicht ausgewählt.

Die Ergänzung des Basis-Indikatoren-Sets durch ausgewählte Indikatoren liefert keine eindeutig besseren Ergebnisse als die wichtigere Benchmark B2: Gemessen an der durchschnittlichen absoluten Abweichung schneidet die Schätzmethode leicht weniger gut ab, gemessen an der durchschnittlichen quadrierten Abweichung hingegen einiges besser. Die Berücksichtigung des Steuerstatus liefert erneut bessere Resultate, welche man jedoch aufgrund des kürzeren Vergleichszeitraums mit einer gewissen Vorsicht interpretieren muss (vgl. Abschnitt 2.5.1).

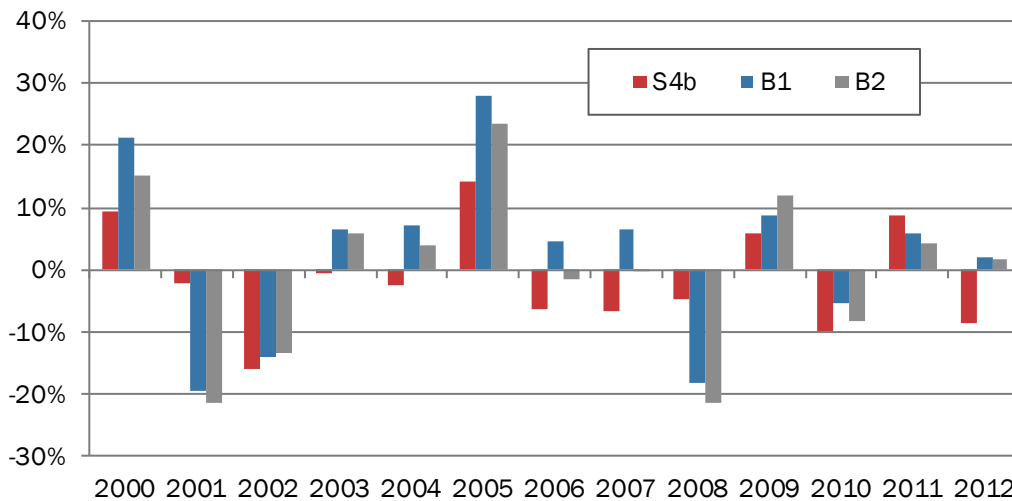
Bei der Schätzmethode S4b wurde untersucht, ob die Prognosequalität erhöht werden kann, wenn für die Jahre, in denen Finanzkrisen stattgefunden haben (2001 und 2008), eine multivariate Schätzung durchgeführt wird. Hierfür wurden beim Indikatoren-Set der Schätzmethode (S4a) etwaige finanzplatzbezogene Indikatoren ausgetauscht, aber dafür das Regressionsmodell um einen finanzplatzbezogenen Interaktionsterm ergänzt. Dieser Interaktionsterm besteht aus der Multiplikation von zwei Variablen: Einer Finanzkrisen-Dummy-Variable (2001 = 1; 2008 = 1; übrige Jahre = 0) und einem Indikatoren-Set bestehend aus der kantonalen Wertschöpfung des Finanzsektors (Aggregat der NOGA08 Branchen 64 bis 66).

Der Grund für diese Vorgehensweise ist, dass in vielen Kantonen die Wertschöpfung des Finanzsektors die bei weitem höchste Korrelation mit den Steuererträgen aufwies. Dennoch wurde der Finanzsektor nur in wenigen Fällen für die Schätzmethode S4a gewählt, da es in Kantonen, in denen der Finanzsektor nur einen geringen Anteil an der Gesamtwertschöpfung hat, nur schwer nachzuvollziehen ist, warum der Finanzsektor zukünftig der beste Indikator für die Gewinnsteuer sein sollte. Der Grund für das hohe Bestimmtheitsmass des Finanzsektors für die Gewinnsteuern in vielen Kantonen liegt darin, dass die Entwicklung der Wertschöpfung der Banken und Versicherungen am besten die Einbrüche der Gewinnsteuern in den Jahren 2001 und 2008 erklärt. Mit der Verwendung der Finanzkrisen-Dummy-Variable können diese beiden Ausreisser-Jahre gut erklärt werden.

Die ex-post Prognosequalität der Schätzmethode S4b ist wesentlich besser als die Schätzmethoden S1-S4a und auch als die beiden Benchmarks. Durch die zusätzliche Desaggregation nach dem Steuerstatus können die Ergebnisse nochmals deutlich verbessert werden. Man muss sich aber bewusst sein, dass der Vergleichszeitraum kürzer ist und die Ergebnisse somit nicht gleich belastbar sind (vgl. Abschnitt 2.5.1). Der Grund, weshalb die Schätzmethode S4b besser abschneidet als die Benchmarks und die zuvor besprochenen Schätzmethoden, besteht darin, dass die Abweichungen von S4b in den Finanzkrisenjahren 2001 und 2008 aufgrund des Finanzkrisenterms in den zugrundeliegenden Regressionen vergleichsweise klein sind, wie in Abb. 2-9 und Abb. 2-8 ersichtlich ist. Wenn man bei der Berechnung der Vergleichsmasse für alle Benchmarks und Schätzungen die Jahre 2001 und 2008 nicht berücksichtigt, dann schneidet die Schätzung S4b nicht mehr besser ab als die Benchmark B2 und

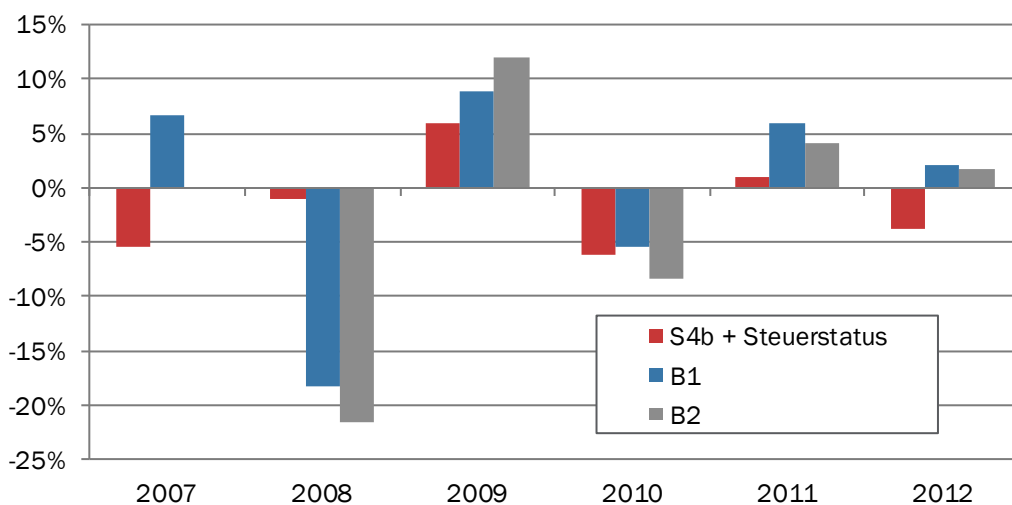
die Schätzungen S1 bis S4a. Die Schätzmethode S4b eignet sich deshalb besser für Sollertragsprognosen, wenn Jahre mit und ohne Finanzkrisen gemischt auftreten; in Zeiten reiner Nicht-Finanzkrisenjahre entsteht kein Vorteil. Hinsichtlich der Verwendung der Schätzmethode im Rahmen von ex-ante Prognosen ergibt sich daraus die Schwierigkeit, dass Finanzkrisen-Jahre bereits bei Erstellung der Prognosen korrekt antizipiert werden müssten.

Abb. 2-8 S4b: Abweichung des geschätzten Steuerertrags pro Jahr



Anmerkung: Abweichung in Prozent des Schätzwertes; ein positiver Wert entspricht einer Unterschätzung des tatsächlichen Steuerertrags, ein negativer Wert einer Überschätzung
Quelle: BAKBASEL

Abb. 2-9 S4b mit Berücksichtigung des Steuerstatus: Abweichung des geschätzten Steuerertrags pro Jahr



Anmerkung: Abweichung in Prozent des Schätzwertes; ein positiver Wert entspricht einer Unterschätzung des tatsächlichen Steuerertrags, ein negativer Wert eine Überschätzung
Quelle: BAKBASEL

2.5.6 Schätzmethode S5

Als Letztes wurde bei der Schätzmethode S5 das Indikatoren-Set mechanisch auf Grundlage des je Kanton bzw. je Steuerstatus höchsten statistischen Bestimmtheitsmasses ausgewählt, ohne auf ökonomisch nachvollziehbare Zusammenhänge und das Vorzeichen der Korrelation zu achten. Wie zu erwarten, wurde mit dieser Schätzmethode die beste ex-post Prognosequalität erreicht. Aufgrund des recht kurzen Beobachtungszeitraum und der grossen Menge an betrachteten Indikatoren dürften jedoch in einigen Kantonen zufällige statistische Zusammenhänge gefunden worden sein. Beispielsweise wurde in einigen Kantonen die Lohnsumme bzw. die Beschäftigung in der Nahrungsmittelindustrie (NOGA 2008-Branche: A10) als bester Indikator für die Entwicklung der Gewinnsteuererträge identifiziert, was ökonomisch schwierig nachzuvollziehen ist. Aus diesem Grund ist anzunehmen, dass die zukünftige Prognosequalität des Indikatoren-Sets S5 geringer ausfallen wird als im Backtest.

2.6 Fazit

Unsere Untersuchung des Zusammenhangs zwischen makroökonomischen Indikatoren und Sollerträgen auf der Basis desaggregierter Daten hat verschiedene Ergebnisse zu Tage gefördert.

Bezüglich der Einkommenssteuern können wir folgende Ergebnisse festhalten:

- Durch die kantonale Desaggregation kann die ex-post Prognosequalität verbessert werden, vgl. insbes. die Schätzmethoden S3, S4a und S4b.
- Eine zusätzliche Desaggregation nach Einkommensklassen scheint hingegen keine zusätzliche Verbesserung der ex-post Prognosequalität zu bringen.
- Es gibt bei der Auswahl der Indikatoren-Sets ein Trade-Off zwischen ihrem theoretischen Appeal und ihrer ex-post Prognosegüte. Je weiter man sich von der aus einer theoretischen Perspektive optimalen Lösung (pendlerbereinigte kantonale Lohneinkommen) hin zu der aus einer statistischen Perspektive optimalen Lösung (numerisch selektiertes Indikatoren-Set auf der Basis des Bestimmtheitsmasses) bewegt, desto besser wird die ex-post Prognosegüte gemessen an den drei von uns verwendeten Massen. Aufgrund des recht kurzen Beobachtungszeitraums und der Vielzahl an getesteten Indikatoren steigt dadurch allerdings die Gefahr von Scheinzusammenhängen, was sich negativ auf die Eignung der Methoden für ex-ante Prognosen auswirkt.

Hinsichtlich der Gewinnsteuern hat unsere Untersuchung Folgendes ergeben:

- Es scheint bei der Gewinnsteuer schwieriger zu sein, allein mit einer kantonalen Desaggregation eine Verbesserung der ex-post Prognosegüte zu erreichen als bei der Einkommenssteuer: Die Ergebnisse zur Schätzmethode S4a sind nicht eindeutig; das heisst, bezüglich zwei der drei Vergleichsmasse konnte eine leichte Verbesserung erreicht werden aber

nicht bezüglich des dritten Masses. Die Schätzmethode S4b, bei welcher ein Finanzkrisenterm in die zugrundeliegenden Regressionen einbezogen wird, bringt zwar eine deutliche Verbesserung der ex-post Prognosegüte, vorrangig für die Finanzkrisenjahre. Eine etwaige Verwendung dieser Methode im Rahmen von ex-ante Prognosen würde jedoch bedingen, dass frühzeitig genug antizipiert werden kann, ob es zu einer Finanzkrise kommt oder nicht. Die Schätzmethode S5 liefert deutliche bessere Ergebnisse als die Benchmarks und alle anderen Methoden, beruht aber auf einer numerischen Selektion der Indikatoren.

- Eine zusätzliche Desaggregation nach Steuerstatus scheint eine Verbesserung der ex-post Prognosegüte zu bringen. Man muss sich aber bewusst sein, dass bei den Schätzungen mit einer Berücksichtigung des Steuerstatus die Vergleichsmasse nur auf der Basis des Zeitraums 2007 bis 2012 anstatt 2000 bis 2012 berechnet werden können, was die Ergebnisse etwas anfälliger für Zufälle und somit auch etwas weniger belastbar macht (vgl. Abschnitt 2.5.1). Auf Basis der wenigen verfügbaren Datenpunkte war eine genauere Untersuchung, ob hier eine bessere Abbildung der Zusammenhänge erreicht werden könnte, nicht möglich.
- Auch bei der Gewinnsteuer scheint es ein Trade-Off zwischen dem theoretischen Appeal der Schätzmethoden und ihrer ex-post Prognosegüte zu geben: Je weiter man sich von der aus einer theoretischen Perspektive optimalen Lösung (kantonales BIP) hin zu der aus einer statistischen Perspektive optimalen Lösung (numerisch selektioniertes Indikatoren-Set auf der Basis des Bestimmtheitsmasses) bewegt, desto besser wird tendenziell die ex-post Prognosegüte. Wie oben schon erwähnt steigt dadurch aber die Gefahr von Scheinzusammenhängen, was sich negativ auf die Eignung der Methoden für ex-ante Prognosen auswirkt.
- Für die Modellierung der Gewinnsteuer könnten ferner Verlustvorträge eine Rolle spielen. Bezüglich dieser Problematik konnten mit der vorhandenen Datenbasis aber keine belastbaren statistischen Ergebnisse erzielt werden. Schätzungen mit zeitlich verzögerten Indikatoren zeigten nur in einzelnen Kantonen einen statistisch signifikanten Zusammenhang mit der Entwicklung der Steuererträge. Falls Verlustvorträge bei der Ergänzung des bisherigen Schätzmodells künftig berücksichtigt werden sollen, sind präzise Daten zu der Entwicklung der Verlustvorträge notwendig.

Zum Abschluss unserer Untersuchung des ersten Zusammenhangs sei hier darauf hingewiesen, dass wir die Prognoseperformance unserer Benchmarks und Schätzmethoden in den Tab. 2-1 und Tab. 2-2 auch noch mit der Performance der ex-ante Sollertragsschätzungen der ESTV verglichen haben, welche uns von der ESTV zur Verfügung gestellt wurden. Bezüglich der Einkommenssteuer lautet das Ergebnis, dass die meisten unserer Schätzmethoden (S3, S4a, S4b, S5) kleinere Abweichungen aufwiesen als die ESTV-Sollertragsschätzungen. Bezüglich der Gewinnsteuer waren die Abweichungen sogar bei allen unseren Schätzmethoden kleiner als bei den ESTV-Sollertragsschätzungen. Man muss sich aber bewusst sein, dass der Performancevergleich zwischen unseren Schätzmethoden und den ESTV-Sollertragsschätzungen aus zwei Gründen nicht wirklich aussagekräftig ist: Erstens handelt es sich bei unseren Schätzungen um ex-post Prognosen, während es sich bei den ESTV-Schätzungen

um ex-ante Prognosen handelt, was die ESTV-Schätzungen gegenüber unseren Schätzungen im Performancevergleich benachteiligt. Zweitens wurden die ESTV-Schätzungen in der Vergangenheit nicht in allen Jahren mit den gleichen Schätzmethoden erstellt (vgl. ESTV 2015), was die Vergleichbarkeit zu unseren Schätzmethoden zusätzlich reduziert.

3 Zusammenhang Sollerträge und Steuereinnahmen

Die Sollerträge eines Bemessungsjahres treffen über mehrere Jahre hinweg als Steuereinnahmen ein. Gemäss ESTV (2015, Seite 9) sind dies erfahrungsgemäss etwa 7 Jahre. Im vorliegenden Kapitel wird dieser zeitliche Zusammenhang zwischen den Sollerträgen und den Steuereinnahmen analysiert und geprüft, ob eine kantonale Desaggregation Potenzial zur Verbesserung der Einnahmenschätzungen bietet. Dieses Kapitel ist folgendermassen aufgebaut: zuerst wird der gewählte Untersuchungsansatz kurz eingeführt, danach auf die verwendeten Daten eingegangen. Anschliessend werden die Ergebnisse ausführlich diskutiert und zum Schluss ein Fazit gezogen.

3.1 Untersuchungsansatz

Bei der Untersuchung des zeitlichen Zusammenhangs zwischen den Sollerträgen und den Steuereinnahmen geht es im Kern um eine Analyse von zeitlichen Verteilungsmustern. Darunter verstehen wir, wie sich die Steuererträge (Sollerträge) eines bestimmten Bemessungsjahres (t_0) auf die Steuereinnahmen des entsprechenden Jahres (t_0) und der nachfolgenden Jahre ($t_1, t_2, t_3...$) aufteilen. Vor diesem Hintergrund kann der Ansatz folgendermassen erläutert werden:

In einem ersten Schritt werden mit deskriptiven statistischen Methoden interkantonale und intertemporale Unterschiede in den zeitlichen Verteilungsmustern untersucht. Mit intertemporalen Unterschieden in den zeitlichen Verteilungsmustern sind Trends im Ablieferungsverhalten von natürlichen und juristischen Personen gemeint. Mit diesem Analyseschritt soll die Basis für die Modellierung im zweiten Analyseschritt geschaffen werden.

In einem zweiten Schritt werden die Steuereinnahmen auf der Basis von (tatsächlichen) Sollerträgen aus der Steuerstatistik und zeitlichen Verteilungsmustern modelliert. Dabei wird insbesondere überprüft, ob die ex-post Prognosequalität der Steuereinnahmen durch eine kantonale desaggregierte Modellierung verbessert werden kann; aufgrund von zuvor identifizierten Trends in den Verteilungsmustern wird in der Modellierung exemplarisch mit zwei verschiedenen Mustern gearbeitet: einem Muster, welches auf der Basis weiter zurückreichender Zeitreihen berechnet wurde und einem Muster, welches auf einem aktuelleren Zeitraum beruht.

3.2 Verwendete Daten

Zur Analyse des zeitlichen Verteilungsmusters lagen uns Zeitreihen der ESTV über die nationalen und kantonalen Steuereinnahmen und den jeweiligen Anteil der Steuererträge der dazugehörigen Bemessungsjahre für die Jahre 1995 bis 2014 vor (ALTO-Files). Für die Analyse des zeitlichen Musters waren keine Daten zur Aufteilung der Steuereinnahmen nach Gewinnsteuer- und Einkommenssteuer-Erträgen vorhanden.

Für die Steuererträge haben wir ergänzend die Daten aus der Steuerstatistik der ESTV hinzugezogen (vgl. Abschnitt 2.2.1). Grundsätzlich lassen sich die Steuererträge

für den Zeitraum 1995 bis 2008¹⁶ über die ALTO Files berechnen. Für den Zeitraum 1995 bis 2002 haben wir auch diese Daten verwendet, um längere Zeitreihen zu erhalten. Für den Zeitraum 2003¹⁷ bis 2012 haben wir dagegen mit den Erträgen aus der Steuerstatistik gearbeitet, da diese als gesicherter gelten, auch wenn die Erträge am aktuellen Rand noch Schätzungen darstellen.

Die Steuererträge eines Bemessungsjahres können noch viele Jahre später als Steuereinnahmen anfallen. Gemäss ESTV (2015, Seite 9) sind dies erfahrungsgemäss etwa 7 Jahre. Die Konsequenz daraus ist, dass uns für die zu berechnenden Verteilungsmuster gesicherte Daten für den Zeitraum 1995 bis 2008 zur Verfügung stehen.¹⁸ Für die deskriptiven Analysen zum kantonalen Vergleich und den zeitlichen Trends greifen wir jedoch auf das Ablieferungsmuster der ersten 4 Jahre (t0, t1, t2, t3) zurück, da in den ersten 4 Jahren ohnehin der Grossteil der Erträge anfällt. Der Vorteil ist, dass dadurch auch aktuellere Entwicklungen bis zum Jahr 2011 dargestellt werden können, auch wenn zu beachten ist, dass die Verteilung der Steuererträge der Jahre 2009 bis 2011 noch auf Schätzungen basiert.

3.3 Unterschiede in den Verteilungsmustern zwischen Kantonen

Betrachtet man die zeitliche Aufteilung der Steuererträge im Zeitraum 1995 bis 2011 auf die Steuereinnahmen der anschliessenden Jahre, so fallen einige Gemeinsamkeiten zwischen den Schweizer Kantonen auf. Im betrachteten Zeitraum wurde in allen Kantonen der bei weitem grösste Teil der Steuerschuld im Jahr t1 nach der Veranlagung entrichtet. In der Schweiz insgesamt lag dieser Anteil in der Periode t1 bei durchschnittlich 79.2 Prozent. In der Periode t2 wurde in der Schweiz (13.0%) und in den meisten Kantonen der zweitgrösste Anteil der Steuerschuld beglichen. In den Perioden t0 und t3 liegt der Ablieferungsanteil in der Schweiz und in den meisten Kantonen unter 5 Prozent (vgl. Abb. 3-1).

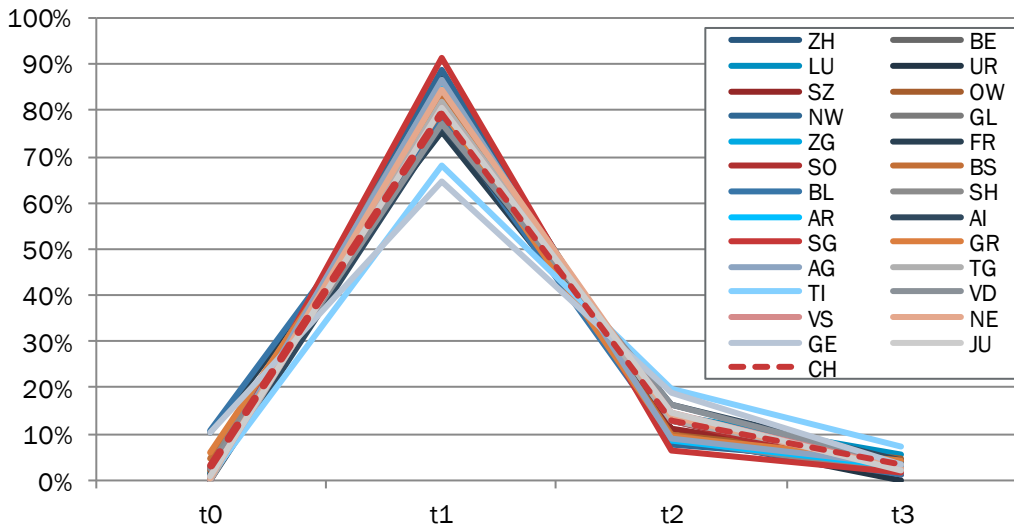
Neben den Gemeinsamkeiten fallen jedoch auch Unterschiede zwischen den Kantonen auf. Die Bandbreite der durchschnittlich beglichenen Steuerschuld in t0 schwankt zwischen 0.0 Prozent (Thurgau, Glarus, St. Gallen und Appenzell Innerrhoden) und 10.6 Prozent (Basel-Land). Im Jahr t1 liegt die Bandbreite zwischen 64.7 (Genf) und 91.5 Prozent (St. Gallen), im Jahr t2 zwischen 6.3 Prozent (St. Gallen) und 19.7 Prozent (Tessin). Erst im Jahr t3 sind die Unterschiede deutlich geringer und einzig im Kanton Tessin liegt der Ablieferungsanteil über der 5-Prozent-Marke. Eine Erklärung für diese teils beträchtlichen Unterschiede ist vor allem bei den unterschiedlichen kantonalen Steuergesetzgebungen sowie in der Veranlagungspraxis zu finden.

¹⁶ Bis 2008 unter der Annahme, dass sich die Steuereinnahmen über einen Zeitraum von 7 Jahren verteilen, was der Erfahrung der ESTV (2015, Seite 9) entspricht.

¹⁷ Ab 2003, weil erst dann alle Kantone bei den Steuern für die natürlichen Personen auf die einjährige Bemessung umgestellt haben.

¹⁸ Von 1995, da dies das erste Jahr ist, für welches wir Steuerertragsdaten haben (ALTO Files); bis 2008, weil von einer 7-jährigen Verzögerungsperiode auszugehen ist.

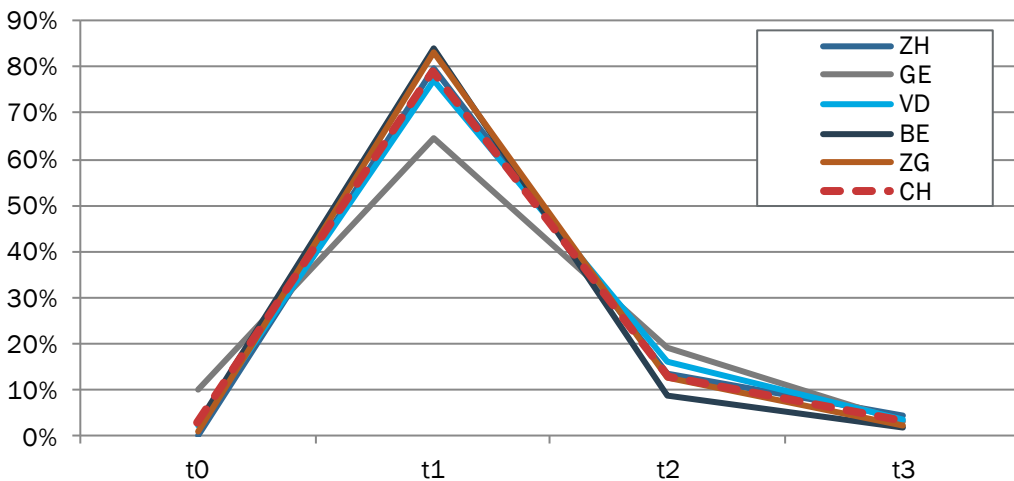
Abb. 3-1 Kantonale Unterschiede in den Verteilungsmustern: Durchschnittliche (1995 - 2011) pro Jahr beglichene Steuerschuld in Prozent



Anmerkung: Durchschnittlich pro Jahr beglichene Steuerschuld in Prozent des Steuerertrags eines Bemessungsjahres für den Zeitraum 1995 - 2011
 Quelle: BAKBASEL, ESTV

Für die Prognose der zeitlichen Ablieferung der Steuererträge in der Schweiz ist vor allem die Entwicklung in den fünf Kantonen mit dem grössten Steueraufkommen von entscheidender Bedeutung (Zürich, Genf, Waadt, Bern und Zug). Zwischen diesen Kantonen fällt das zeitliche Verteilungsmuster etwas homogener als im Vergleich aller 26 Kantone aus. Der Kanton Genf weicht jedoch deutlich von den anderen grossen Steuerzahler-Kantonen ab (vgl. Abb. 3-2).

Abb. 3-2 Kantonale Unterschiede in den Verteilungsmustern bei den 5 grössten Steuerzahler-Kantonen: Durchschnittliche (1995 - 2011) pro Jahr beglichene Steuerschuld in Prozent

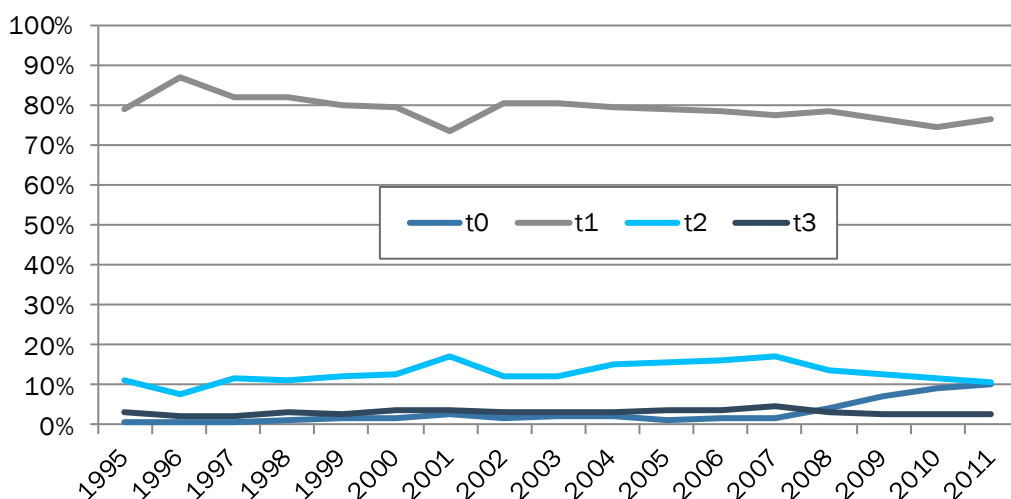


Anmerkung: Durchschnittlich in den Perioden t0 - t3 beglichene Steuerschuld in Prozent des Steuerertrags eines Bemessungsjahres für den Zeitraum 1995 - 2011 in den Kantonen Zürich, Genf, Waadt, Bern, Zug und der Schweiz
 Quelle: BAKBASEL, ESTV

3.4 Unterschiede in den Verteilungsmustern über die Zeit

Im Zeitraum 1995 bis 2011 war in der Schweiz eine deutliche Veränderung des Ablieferungsverhaltens zu verzeichnen. Der Anteil der Steuerschuld, welche bereits im Jahr t0 beglichen wird, ist von 0.7 Prozent im Jahr 1995 auf mehr als 10 Prozent im Jahr 2011 gestiegen. Vor allem seit 2008 war ein starker Anstieg des Anteils zu beobachten. Einschränkend muss jedoch erwähnt werden, dass der Steuerertrag für die Jahre 2010 und 2011 noch nicht gesichert ist und sich daher auch die Anteilswerte für diese Jahre noch ändern können. Während der Anteil der beglichenen Steuerschuld in der Periode t0 im Zeitverlauf gestiegen ist, sank der Anteil im Jahr t1 im Schweizer Durchschnitt (von 79.3% im Jahr 1995 auf 76.6% im Jahr 2011). In der Periode t2 lag der Anteil sowohl 1995 als auch 2011 bei etwas mehr als 10 Prozent, in den Zwischenjahren gab es jedoch kräftigere Schwankungen. Das Ablieferungsverhalten in Periode t3 blieb zwischen 1995 und 2011 relativ konstant.

Abb. 3-3 Zeitliche Veränderung in den Verteilungsmustern: Prozentualer Anteil der beglichenen gesamtschweizerischen Steuerschuld in den Jahren t0 bis t3 über die Zeit



Anmerkung: Anteil der in den Perioden t0-t3 beglichenen Steuerschuld am gesamten Steuerertrag eines Bemessungsjahres in Prozent für den Zeitraum 1995 – 2011
Quelle: BAKBASEL, ESTV

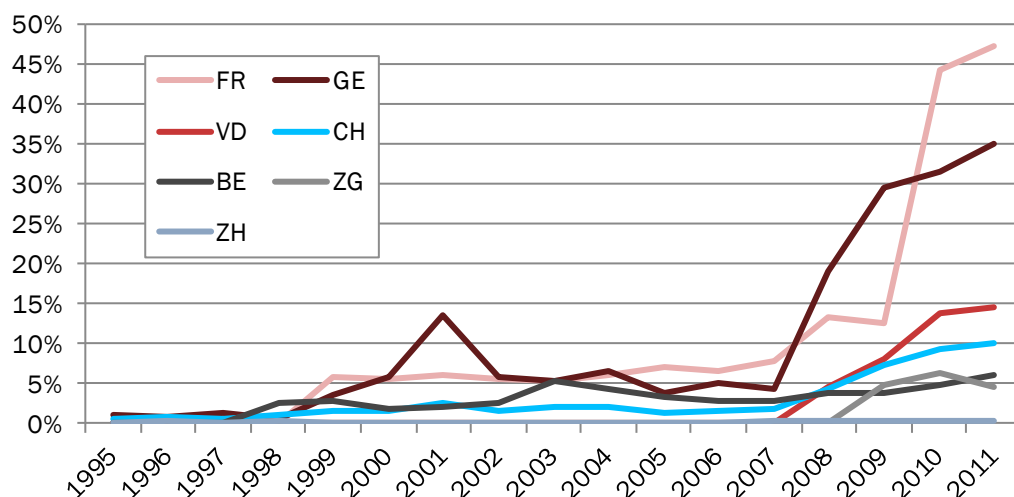
Die Veränderung des zeitlichen Musters betrifft jedoch nicht alle Kantone gleichermaßen. In Zürich ist z.B. der Anteil der in Periode t0 beglichenen Steuerschuld zwischen 1995 und 2011 nur geringfügig von 0.2 auf 0.5 Prozent gestiegen, dafür gab es jedoch grössere Schwankungen beim Anteil der Periode t2. Im Kanton Genf ist der Anteil der Periode t0 dagegen zwischen 1995 und 2011 von 1.1 Prozent auf 35.2 Prozent gestiegen. Auch in den Perioden t1-t3 zeigt sich im Kanton Genf eine sehr hohe Schwankungsbreite bezüglich des Ablieferungsanteils.

Im Folgenden wird kurz auf verschiedene mögliche Faktoren eingegangen, welche hinter den Trends in den zeitlichen Verteilungsmustern stehen könnten.

Ein erster möglicher Grund für die Veränderung des zeitlichen Zusammenhangs zwischen Steuererträgen und Steuereinnahmen in der Schweiz können Änderungen der

Steuergesetzgebung bzw. Änderungen im Ablieferungsverhalten der Kantone sein. Ein interessantes Beispiel für eine solche Entwicklung ist, dass verschiedene Kantone in den letzten Jahren die Bewilligung erhalten haben, Rechnungen für die Vorauszahlung von Steuerschulden der direkten Bundessteuer in Periode t0 zu verschicken. Konkret ist dies seit 2008 in den Kantonen Genf und Waadt, seit 2010 im Kanton Freiburg und seit kürzerem auch im Kanton Jura der Fall. Obwohl die Steuersubjekte selber entscheiden können, ob sie diese Rechnungen in der Periode t0 oder zu einem späteren Zeitpunkt begleichen wollen, liegt die Vermutung nahe, dass diese Praxis die Vorauszahlung fördert und der Anteil von t0 deshalb ansteigt. Die nachfolgende Abbildung legt nahe, dass diese Vermutung korrekt ist: In Genf und Waadt steigt der Anteil von t0 in 2008 sehr deutlich an, in Freiburg ist das Gleiche in 2010 der Fall, was auch den gesamtschweizerischen Durchschnitt für t0 nach oben zieht, weil es sich bei Genf und Waadt um zwei wichtige Steuerkantone handelt. In den Kantonen Bern, Zürich und Zug, in welchen keine entsprechende Rechnungsstellungspraxis eingeführt wurde, steigt der Anteil der in t0 beglichene Steuerschuld nicht oder zumindest unterdurchschnittlich stark an.

Abb. 3-4 Zeitliche Veränderungen in den Verteilungsmustern: Entwicklung Ablieferungsanteils von t0 über die Zeit in ausgewählten Kantonen und der Schweiz



Anmerkung: Durchschnittlich in den Perioden t0 beglichene Steuerschuld in Prozent des Steuerertrags eines Bemessungsjahres für den Zeitraum 1995 – 2011 in den Kantonen Freiburg, Waadt, Bern, Zürich, Genf, Zug und der Schweiz
Quelle: BAKBASEL, ESTV

Ein zweiter Erklärungsansatz für diese Veränderung des zeitlichen Musters ist der starke Rückgang des Zinsniveaus seit der Finanzkrise 2008. Da viele Kantone Vorauszahlungszinsen für Steuerzahlungen vor Fälligkeit anbieten, ist aufgrund des gegenwärtigen Tiefzinsumfeldes die Attraktivität von Steuerzahlungen in der Periode t0 in den letzten Jahren deutlich gestiegen. Dies könnte eben zu der beobachteten Verschiebung im Zahlungsmuster geführt haben, mit entsprechenden Rückwirkungen auf die Steuereinnahmen. Eine einfache Korrelationsanalyse zwischen der Rendite 10-jähriger Schweizer Staatsanleihen (als Proxy für das Zinsniveau) und dem Anteil der Ablieferungen in t0 über den Zeitraum 1995 bis 2011 stützt diese Hypothese (Korrelationskoeffizient -0.74). Wenn diese Hypothese korrekt ist, dürfte angesichts

der anhaltend tiefen Zinsen auch im Zeitraum 2013 bis 2015 der Anteil der Steuerzahlungen in der Periode t_0 hoch ausgefallen sein, was voraussichtlich auch noch einige Zeit anhalten dürfte. Unklar ist allerdings, ob auch der Trend zu immer höheren Zahlungen in t_0 noch weiter anhält, insbesondere weil die Guthabenverzinsung für vorzeitige Steuerzahlungen inzwischen stark gesenkt wurde. Von besonderer Relevanz wäre dies aber, sobald die Zinsen wieder steigen: wenn sich die Situation mittelfristig wieder normalisiert und sich die Anreize bezüglich frühzeitiger Steuerzahlungen damit wieder umkehren, dürfte auch der Anteil von t_0 an den Steuerablieferungen wieder sinken. Dies könnte während der Übergangsphase zu empfindlich tieferen Steuereinnahmen führen.

Ein weiterer möglicher Erklärungsansatz für Schwankungen des Ablieferungsmusters sind die Konjunkturzyklen. Da die Höhe von Steuervorauszahlungen vom vergangenen Einkommen bzw. Gewinn abhängt,¹⁹ kann ein Konjunkturabschwung, der zu einem Rückgang der Einkommen bzw. Gewinne führt, dazu beitragen, dass der Ablieferungsanteil von t_0 zunimmt, da die letztendlich zu zahlende Steuerschuld sinkt. Dies kann unter Umständen, je nach genauem Muster, auch für t_1 entsprechend gelten. Dies könnte nach der Finanzkrise 2008 ebenfalls zum Anstieg des Ablieferungsanteils von t_0 in den letzten Jahren beigetragen haben. Ein starker Konjunkturaufschwung dürfte dementsprechend eine dämpfende Wirkung auf den Ablieferungsanteil in t_0 haben. Wenn auch von den Ursachen her ein ganz anderer Punkt, so könnte sich dies jedoch ähnlich wie der vorangegangene Punkt auf die zukünftigen Steuereinnahmen auswirken. Leider gilt auch hier, dass der Umfang und Detaillierungsgrad der verfügbaren Daten nicht ausreichen, um mit empirischen Methoden diese Hypothese mit einer ausreichenden Gewissheit bestätigen oder verwerfen zu können.

3.5 Methodologische Vorgehensweise bei der Modellierung

In einem nächsten Analyseschritt haben wir die Steuereinnahmen der Schweiz auf der Basis von desaggregierten Daten (d.h. desaggregierten Sollerträgen aus der Steuerstatistik und kantonalen zeitlichen Verteilungsmustern) modelliert. Um die ex-post Prognosequalität der so geschätzten Zeitreihen beurteilen zu können, haben wir sie mit der Performance von zwei Benchmarks verglichen. In diesem Abschnitt wird zunächst die Berechnung der beiden Benchmarks erläutert und danach unsere Schätzmethoden.

Benchmarks

Die von der ESTV im Schätzmodell verwendeten Verteilungsmuster sind nicht fixiert, das heisst, die ESTV überprüft die im Schätzmodell verwendeten zeitlichen Verteilungsmuster laufend und nimmt bei Bedarf Anpassungen vor. In ESTV (2015) werden zwei im aktuellen Schätzmodell verwendete Verteilungsmuster spezifiziert, eines für die natürlichen Personen und eines für die juristischen Personen. Auf der Basis dieser beiden Muster haben wir zwei Zeitreihen zu den Schweizer Steuereinnahmen berechnet, welche uns als Benchmarks zur Beurteilung der ex-post Prognosegüte der von uns geschätzten Reihen dienen. Beiden Benchmarks ist gemein, dass wie im gegenwärtigen Schätzmodell der ESTV keine Desaggregation nach Kantonen vorgenommen wird.

¹⁹ Bzw. die Vorauszahlungen von den bisherigen Steuerzahlungen (adaptive Erwartungen), was durch provisorische Steuerrechnungen, die auf Steuererklärungen früherer Jahre basieren, noch verstärkt werden kann.

Bei der Berechnung der Benchmark B1 sind wir folgendermassen vorgegangen. Zunächst haben wir auf der Basis des Verteilungsmusters für die natürlichen Personen und den Sollerträgen für die natürlichen Personen aus der Steuerstatistik (2003 bis 2012) die Steuereinnahmen für die natürlichen Personen in den Jahren 2009-2012 berechnet. Analog dazu haben wir eine Zeitreihe für die Einnahmen der juristischen Personen berechnet. Diese beiden Zeitreihen haben wir dann zu einer Reihe für die totalen Schweizer Steuereinnahmen (juristische und natürliche Personen) für den Zeitraum 2009 bis 2012 aggregiert.

Bei der Benchmark B2 haben wir hingegen mit einem durchschnittlichen Verteilungsmuster für natürliche und juristische Personen gearbeitet, indem wir das ungewichtete arithmetische Mittel der beiden oben beschriebenen Verteilungsmuster aus ESTV (2015) gebildet haben. Der Vorteil bei dieser Benchmark ist, dass wir mit diesem durchschnittlichen Verteilungsmuster auch weiter zurückliegende Einnahmen berechnen können, weil die Erträge vor 2003 in den ALTO Files nur für natürliche und juristische Personen zusammen vorliegen. Dies erlaubt einen Vergleich der Ergebnisse unserer Schätzung mit der Benchmark B2 für die Jahre 2001 bis 2012.

Schätzmethoden

Insgesamt haben wir vier verschiedene Schätzmethoden verwendet: zwei davon beruhen auf nationalen Daten, zwei auf kantonalen Daten. Dabei sind wir folgendermassen vorgegangen.

Zunächst wurde für jeden Kanton und die Schweiz auf Basis der historischen Steuerdaten das zeitliche Muster der Ablieferung der Steuererträge berechnet. Für das zeitliche Muster wurden die Durchschnittswerte der Jahre 1995 bis 2008 (Variante 1) sowie 2003 bis 2008 (Variante 2) verwendet. Der Vorteil der ersten Variante ist, dass man mehr Datenpunkte für die Durchschnittsbildung zur Verfügung hat. Der Vorteil der zweiten Variante ist hingegen, dass die verwendeten Datenpunkte aktueller sind und deshalb aufgrund der oben diskutierten Trends in den zeitlichen Mustern das heutige, noch unbekannte Muster besser abbilden sollten.²⁰

Als nächstes wurden die erhaltenen Durchschnittswerte für das Verteilungsmuster der Perioden t_0 bis t_7 mit den historischen kantonalen bzw. nationalen Steuererträgen multipliziert. Das Ergebnis waren geschätzte Zeitreihen für die kantonalen bzw. Schweizer Steuereinnahmen. Die kantonalen Zeitreihen wurden aufsummiert, um eine zweite Zeitreihe für die Schweizer Steuereinnahmen zu erhalten (desaggregierter Ansatz). Insgesamt haben wir so vier geschätzte Zeitreihen (S1-S4) für die Schweizer Steuereinnahmen berechnet: einmal nach Kantonen desaggregiert und einmal nicht nach Kantonen desaggregiert; einmal für Verteilungsmuster nach Variante 1 und einmal für das Muster nach Variante 2.

Vergleichsmasse

Die vier von uns geschätzten Zeitreihen der Schweizer Steuereinnahmen und die ESTV Benchmark wurden nun im Backtest auf ihre ex-post Prognosequalität getestet. Hierzu wurden die geschätzten Einnahmen mit den historischen Werten verglichen

²⁰ Allerdings sind aufgrund der noch nicht gesicherten Daten genau diejenigen Jahre mit den grössten Verschiebungen im zeitlichen Muster (siehe deskriptive Analyse) noch in keiner der beiden Varianten enthalten.

und die Prognosequalität mit den gleichen drei Massen wie in Kapitel 2 verglichen (vgl. Abschnitt 2.3.5 für die Definition der Masse).

3.6 Ergebnisse der Modellierung

Die Tabelle 3-1 fasst unsere Resultate zur Modellierung des zeitlichen Zusammenhangs zwischen Sollerträgen und Steuereinnahmen zusammen. Im Wesentlichen wurden drei Ergebnisse erreicht, welche im Folgenden erläutert werden:

Tab. 3-1 Vergleich der ex-post Prognosequalität von verschiedenen Benchmarks und Schätzmethoden für die Steuereinnahmen

Benchmarks (B) bzw. Schätzungen (S)		Durchschnitt der absoluten Abweichungen pro Jahr		Durchschnitt der quadrierten Abweichungen pro Jahr		Bestimmtheitsmass
		Abweichung in Franken	% Veränderung ggü. Abweichung B2	Abweichung in Franken	% Veränderung ggü. Abweichung B2	
B1	Ohne kantonale Desaggregation, auf Basis der Verteilungsmuster in ESTV (2015), NP und JP getrennt behandelt	(1.08E+09)	(79%)	(1.26E+18)	(116%)	(0.57)
B2	Ohne kantonale Desaggregation, auf Basis der Verteilungsmuster in ESTV (2015), NP und JP nicht getrennt behandelt	6.04E+08	0%	5.83E+17	0%	0.34
S1	Ohne kantonale Desaggregation, auf Basis des durchschnittlichen Verteilungsmusters 1995-2008	5.38E+08	-11%	4.78E+17	-18%	0.35
S2	Mit kantonalen Desaggregation, auf Basis des durchschnittlichen Verteilungsmusters 1995-2008	5.31E+08	-12%	4.74E+17	-19%	0.37
S3	Ohne kantonale Desaggregation, auf Basis des durchschnittlichen Verteilungsmusters 2003-2008	4.68E+08	-23%	3.29E+17	-44%	0.35
S4	Mit kantonalen Desaggregation, auf Basis des durchschnittlichen Verteilungsmusters 2003-2008	4.57E+08	-24%	3.19E+17	-45%	0.37

Bemerkungen: Bei B2 und S1 bis S4 wurden die Steuereinnahmen der natürlichen und juristischen Personen aufgrund der Datenlage nicht getrennt behandelt. Der Vergleichszeitraum umfasst bei der Benchmark B1 die Jahre 2009 bis 2012 (Werte in runden Klammern), bei der Benchmark B2 und den Schätzungen S1 bis S4 hingegen die Jahre 2001 bis 2012 (Werte ohne runde Klammern); die Ergebnisse für Benchmark B1 sind deshalb nicht ohne Weiteres mit den anderen Ergebnissen vergleichbar, weil die Abweichungen seit 2009 tendenziell gross waren (betrifft alle drei Vergleichsmasse) und weil das Bestimmtheitsmass mit einer Verkürzung der Zeitreihen ansteigt (betrifft nur das Bestimmtheitsmass). Die Spalten mit der Bezeichnung „Abweichung in Franken“ geben die durchschnittliche jährliche (absolute bzw. quadrierte) Abweichung der geschätzten Werte von den tatsächlichen Werten in Franken an (vgl. Text für die Formel); die dabei verwendete E-Notation kann so erläutert werden: 1.00E+09 = 1.00 * 10⁹ = 1 Mia. Die Spalten mit der Bezeichnung „% Veränderung ggü. Abweichung B2“ geben an, wie sich der entsprechende Prognosefehler relativ zum Prognosefehler der Benchmark B2 verändert; das heisst, um wieviel Prozent die durchschnittliche jährliche (absolute bzw. quadrierte) Abweichung in Franken gegenüber der Abweichung bei Benchmark B2 reduziert/verbessert (negative Prozentzahlen) bzw. gesteigert/verschlechtert wird (positive Prozentzahlen).
Quelle: BAKBASEL

Ein erstes Ergebnis lautet, dass alle von uns geschätzten Zeitreihen bessere Prognoseergebnisse als die beiden Benchmarks liefern. Man muss hier anmerken, dass die Werte für die Benchmark B1 nur für den Zeitraum von 2009 bis 2012 vorliegen. Da die Abweichungen in diesen 4 Jahren bei allen Schätzungen verhältnismässig gross sind, sind die Tabellenwerte der Benchmark B1 (in runden Klammern) nicht direkt mit den Tabellenwerten der Benchmark B2 und der Schätzmethoden (Werte ohne

runde Klammern) vergleichbar, welche auf dem Vergleichszeitraum 2001 bis 2012 beruhen. Ein Blick auf die nachfolgende Abbildung macht aber ohne weitere Berechnungen deutlich, dass die obige Aussage nicht relativiert werden muss: die Schätzmethoden S1 bis S4 schneiden auch im Vergleichszeitraum 2009 bis 2012 besser ab als beide Benchmarks.

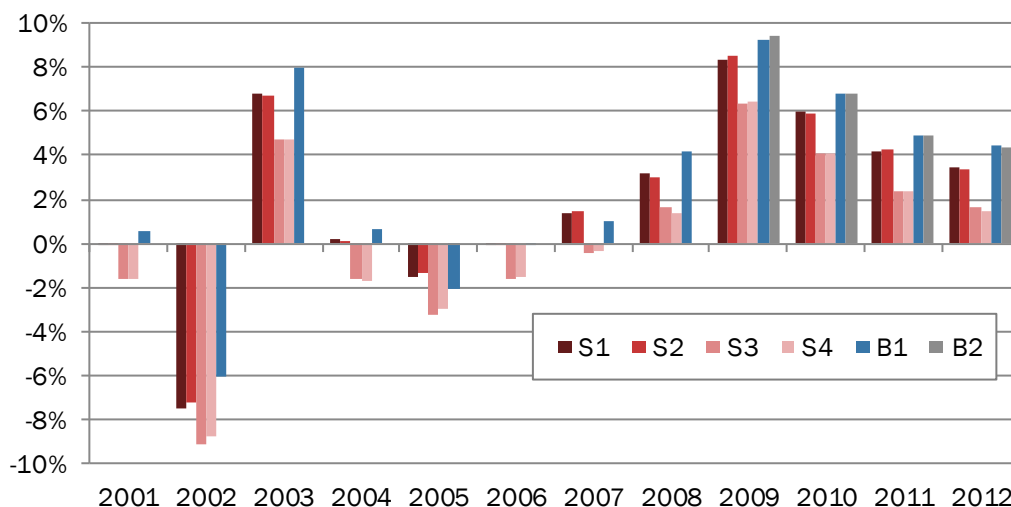
Eine mögliche Erklärung für die bessere ex-post Prognoseperformance könnte darin bestehen, dass unsere Schätzungen auf Verteilungsmustern basieren, welche für ex-post Prognosen optimiert sind, während die Benchmarks auf Verteilungsmustern des aktuellen Schätzmodells basieren, welche für ex-ante Prognosen optimiert sind (d.h. möglichst gute Einnahmenschätzungen in den Jahren nach 2012 bringen sollten). Unsere nachfolgenden Ergebnisse sind aber unabhängig von dieser Frage, weil sie auch alleine auf der Basis eines Vergleichs der ex-post Prognosequalität unserer vier Schätzmethoden gewonnen werden können.

Ein zweites wichtiges Ergebnis lautet, dass die nach Kantonen desaggregierten Schätzmethoden (bzw. die Ausnutzung der Kantonsvarianz) jeweils leichte Verbesserungen gegenüber der nicht-desaggregierten Schätzmethoden (bzw. Schätzung auf der Ebene Schweiz) bringen, wie der Vergleich von S1 mit S2 und von S3 mit S4 deutlich macht. Aufgrund der im Abschnitt 3.3 diskutierten Kantonsvarianz ist dieses Ergebnis von der Richtung her nicht unerwartet, fällt aber aufgrund der beträchtlichen Kantonsvarianz deutlich kleiner als erwartet aus.

Als drittes Ergebnis können wir festhalten, dass sich die Prognosequalität verbessert, wenn man bei der Bildung der durchschnittlichen zeitlichen Verteilungsmuster nur die Jahre 2003 bis 2008 berücksichtigt, anstelle des weiter zurückreichenden Zeitraums 1995 bis 2008. Dies deutet darauf hin, dass die Wahl des Zeitraums, auf dessen Basis man die Verteilungsmuster berechnet, eine wichtige Rolle spielt. Soweit die verfügbaren Daten Aussagen zulassen, scheint sich hier also eine Veränderung des zeitlichen Ablieferungsmusters zu zeigen, und zwar bereits vor den oben genannten spezifischen institutionellen Änderungen (vgl. Kap. 3.4). Denkbar ist jedoch ein Zusammenhang mit dem Übergang zur einjährigen Besteuerung, der 2003 in allen Kantonen abgeschlossen war.

Das dritte Ergebnis zeigt sich auch in der nachfolgenden Abbildung, welche die jährlichen Abweichungen für die 4 Schätzmethoden sowie für die beiden Benchmarks in Prozent des Schätzwertes graphisch darstellt. Zu Beginn waren die auf dem weiter zurückreichenden Zeitraum (1995-2008) basierenden Schätzungen S1 und S2 noch besser als die Schätzungen S3 und S4; seit 2007 sind aber die auf dem aktuelleren Zeitraum (2003-2008) basierenden Schätzungen S3 und S4 besser.

Abb. 3-5 Abweichung der geschätzten Steuereinnahmen pro Jahr



Anmerkung: Abweichung in Prozent des Schätzwertes; ein positiver Wert entspricht einer Unterschätzung der tatsächlichen Steuereinnahmen, ein negativer Wert einer Überschätzung.

Quelle: BAKBASEL

3.7 Fazit

Unsere Untersuchung des zeitlichen Zusammenhangs zwischen Sollerträgen und Steuereinnahmen hat verschiedene Ergebnisse zu Tage gefördert:

- Es gibt beträchtliche kantonale und intertemporale Variationen in den Verteilungsmustern.
- Eine Modellierung des Zusammenhangs auf Basis einer kantonalen Desaggregation scheint trotz der grossen kantonalen Unterschiede nur eine leichte Verbesserung der ex-post Prognosegüte zu bewirken.
- Die Wahl des Zeitraums, auf dessen Basis man die zeitlichen Verteilungsmuster berechnet, scheint einen grösseren Einfluss auf die ex-post Prognosequalität zu haben als die kantonale Desaggregation. Man muss sich aber bewusst sein, dass die ESTV diesem Umstand bereits insofern Rechnung trägt, als die im Schätzmodell verwendeten Verteilungsmuster laufend überprüft und angepasst werden. Die erheblichen und nicht in allen Kantonen gleich verlaufenden Veränderungen der Verteilungsmuster über die Zeit (gerade auch in den letzten verfügbaren Jahren) deuten jedoch darauf hin, dass hier noch weiteres Verbesserungspotential schlummern könnte, insbesondere wenn es gelingen würde, die Ursachen für die Veränderungen der Verteilungsmuster in die Modellierung mit einzubeziehen (z.B. konjunkturelle Einflüsse, regulatorische Veränderungen).

4 Schlussfolgerungen und Empfehlungen

Das Ziel dieses Gutachtens bestand darin, zwei dem bisherigen Schätzmodell der direkten Bundessteuer zugrundeliegende Zusammenhänge auf der Basis von desaggregierten Daten in einer ex-post Analyse zu untersuchen, um herauszufinden, ob in einer desaggregierten Modellierung Potenzial für eine Verbesserung der Prognosequalität liegt. Konkret ging es dabei um die folgende Zusammenhänge: erstens, um den Zusammenhang zwischen makroökonomischen Indikatoren und Sollerträgen der Einkommens- bzw. Gewinnsteuer; zweitens, um den (zeitlichen) Zusammenhang zwischen den Sollerträgen und den Steuereinnahmen.

Die Ergebnisse der Kapitel 2 und 3 deuten insgesamt darauf hin, dass es bei beiden untersuchten Zusammenhängen ein gewisses Verbesserungspotenzial hinsichtlich der ex-post Prognosequalität gibt. Das Potential für mögliche Verbesserungen ist in den einzelnen Schritten zwar meistens eher klein bis vernachlässigbar, jedoch ist in der Summe der verschiedenen Punkte zumindest bezüglich der ex-post Prognosequalität ein spürbarer Effekt ermittelt worden.

Im Einzelnen scheint das Potenzial für eine Verbesserung der ex-post Prognosequalität durch eine desaggregierte Modellierung beim Zusammenhang zwischen makroökonomischen Indikatoren und Sollerträgen grösser als beim zeitlichen Zusammenhang zwischen den Sollerträgen und den Steuereinnahmen; bei Letzterem scheint die Erfassung oder Modellierung von zeitlichen Trends in den Verteilungsmustern einen grösseren Einfluss zu haben.

Noch keine abschliessenden Aussagen sind damit zum Potential einer Verbesserung der ex-ante Prognosequalität möglich. Die Analysen beruhen auf ex-post Untersuchungen. Nicht jede Verbesserung der ex-post (oder in-sample) Prognosegüte kann auch in eine Verbesserung der ex-ante (oder out-of-sample) Prognosegüte für die Zukunft überführt werden. Es ist jedoch auch möglich, dass die positiven Effekt ex-ante sogar grösser ausfallen als ex-post. Im vorliegenden Fall ist allerdings aus verschiedenen Gründen zu vermuten, dass das ex-ante Verbesserungspotential eher kleiner ausfällt als das ex-post Verbesserungspotenzial. Beispielsweise können gute Ergebnisse für bestimmte Indikatoren angesichts des sehr kurzen Beobachtungszeitraums auf nur temporär geltenden Zusammenhängen basieren. Selbst wenn dies nicht der Fall ist, so ist zum einen die Fehlerrate ex-ante typischerweise grösser als ex-post, sodass eine Verbesserung im gleichen Umfang den Fehler prozentual weniger reduziert. Zum anderen sind diejenigen Kenngrössen, die in einer desaggregierten ex-ante Prognose Verwendung finden, in der Regel selbst mit mehr Unsicherheiten behaftet als Kennzahlen auf einer höheren Aggregationsebene. Die Fragestellung nach potenziellen ex-ante Verbesserungspotenzial war jedoch nicht Teil des Analyseauftrags, womit diese Aussagen zur Relevanz der Ergebnisse für die ex-ante Prognosequalität bisher nur den Charakter von Hypothesen haben. In jedem Fall zeigt die Studie Ansatzpunkte auf, die ein gewisses Potenzial haben könnten, die bisherigen Schätzmodelle zu den Steuererträgen und -einnahmen zu optimieren.

Zusammenhang zwischen Indikatoren und Sollerträgen

Bei der Prognose der Einkommens-Steuererträge erscheint der desaggregierte Ansatz auf Kantonsebene grundsätzlich als prüfungswerte Alternative. Allerdings reicht die Verwendung des theoretisch sinnvollsten Indikators – der kantonalen pendlerbereinigten Lohnsummen – nicht aus, um die ex-post Prognosegüte zu erhöhen. Erst bei der Verwendung eines kantonsspezifischen Indikatoren-Sets aus mehreren ökonomischen Grössen (Primäreinkommen CH, Lohnsumme CH und Kantone, Nominales BIP CH und Kantone sowie Bruttowertschöpfung CH und Kantone) kann die ex-post Schätzgenauigkeit verbessert werden. Dies zeigt, dass die Einkommensteuer-Entwicklung in den Kantonen unterschiedlich verläuft und von abweichenden ökonomischen Erklärungsfaktoren bestimmt wird. Durch die Verwendung von weiteren ökonomischen Indikatoren (z.B. Lohnsumme Finanzsektor in einigen Kantonen) kann die ex-post Prognosequalität noch weiter verbessert werden, allerdings steigt dadurch auch die Gefahr von zufälligen statistischen Korrelationen aufgrund der recht kurzen Beobachtungsperiode. Insgesamt erachten wir die Desaggregation nach Kantonen als eine prüfungswerte Erweiterung der bisherigen Schätzmethoden für die Einkommenssteuer. Bevor aber eine Erweiterung der bisherigen Schätzmethoden in Betracht gezogen würde, müssten in einem nächsten Schritt die besten desaggregierten Schätzmethoden und die ihnen zugrundeliegenden Indikatoren noch auf ihre ex-ante Prognosequalität untersucht werden.

Die zusätzliche Desaggregation nach Einkommensklassen scheint dagegen – entgegen der ursprünglichen Hypothese – keine weitere Verbesserung der ex-post Prognosegüte zu bringen. Sie ist daher aufgrund des steigenden Aufwandes zumindest bei der gegenwärtigen Datenlage nicht zu empfehlen.

Bei den Gewinnsteuererträgen war es relativ schwierig, die ex-post Prognosequalität alleine auf der Basis einer kantonalen Desaggregation zu verbessern. Die Schätzung der kantonalen Steuererträge mit einem kantonalen Indikatoren-Set hat die Prognosequalität im Backtest gegenüber den bisherigen Schätzmethoden nicht verbessert. Eine deutliche Verbesserung der Prognosequalität wird erst erreicht, wenn das Regressionsmodell um einen finanzplatzbezogenen Interaktionsterm für die beiden Finanzkrisen 2001 und 2008 ergänzt wird. Eine etwaige Verwendung dieser Methode im Rahmen von ex-ante Prognosen würde jedoch bedingen, dass frühzeitig genug antizipiert werden kann, ob es zu einer Finanzkrise kommt oder nicht.

Die Unterteilung nach Steuerstatus der Unternehmen als zusätzliche Desaggregationsstufe verbessert die ex-post Prognosequalität bei allen getesteten Methoden. Man muss hier aber berücksichtigen, dass der Zeitraum für die Beurteilung der ex-post Prognosequalität aufgrund der Datenlage bei den Schätzungen mit einer zusätzlichen Desaggregation nach Steuerstatus kleiner ist als bei den Schätzungen ohne eine solche Desaggregation, was die Ergebnisse weniger belastbar macht. Trotzdem können wir festhalten, dass unsere Ergebnisse zur Desaggregation nach Steuerstatus auf ein gewisses Verbesserungspotenzial hindeuten. Dies möglicherweise selbst dann, wenn der Steuerstatus im Zuge der Unternehmenssteuerreform III abgeschafft würde. Denn der Unternehmensstatus kann nicht nur selbst die Entwicklung der Steuerbeträge beeinflussen. Vielmehr erscheint es plausibel, den Steuerstatus als Proxy-Variable für andere Indikatoren, die nicht direkt beobachtet werden können, zu interpretieren. Mögliche Indikatoren sind die Branchenzugehörigkeit, die Internationalität der Gesellschaft (z.B. Exportlastigkeit), die Unternehmensgrösse oder weitere

Unternehmensmerkmale. Diese Indikatoren können bisher mangels entsprechender desaggregierter Daten nicht direkt untersucht bzw. genutzt werden. Wir erachten es deshalb als prüfenswert, Informationen wie Branchenzugehörigkeit, Unternehmensgrösse, Internationalität etc. in Zukunft systematisch zu erheben und in einigen Jahren auf das mit ihnen möglicherweise verbundene Verbesserungspotenzial zu überprüfen.

Sowohl bei den Gewinn- als auch den Einkommensteuererträgen könnte es prüfenswert sein, eine Mischung aus mehreren Schätzmethoden zu verwenden, um das derzeitige Schätzmodell zu ergänzen. Eine mögliche Vorgehensweise wäre z.B. die zukünftigen Erträge jeweils mit dem bisherigen Schätzmodell sowie einer oder mehrerer desaggregierten Schätzmethoden zu schätzen. Der finale Prognosewert würde sich dann als (gewichteter) Durchschnitt dieser Methoden ergeben. Ein Nachteil dieses Vorgehens ist jedoch, dass die Komplexität und der Arbeitsaufwand für die Prognose der Steuererträge ansteigen würden.

Zeitlicher Zusammenhang zwischen Sollerträgen und Steuereinnahmen

Das zweite Ziel der vorliegenden Studie war es, zu testen, ob ein desaggregierter kantonaler Ansatz Vorteile bei der Schätzung des zeitlichen Zusammenhangs zwischen Steuererträgen und Steuereinnahmen liefert. Dabei konnte gezeigt werden, dass ein solcher desaggregierter Ansatz die ex-post Prognosegüte leicht verbessert, was aufgrund der relativ starken Kantonsvarianz bei den Verteilungsmustern kein unerwartetes Ergebnis darstellt.

Es hat sich aber gleichzeitig gezeigt, dass es im Vergleich zur Frage, ob man die Kantonsvarianz berücksichtigen sollte oder nicht, eigentlich eine wichtigere Rolle spielt, welchen Zeitraum man zur Berechnung der durchschnittlichen Verteilungsmuster wählt. Die ESTV trägt diesem Umstand bereits insofern Rechnung, als die im Schätzmodell verwendeten Verteilungsmuster laufend überprüft und angepasst werden. Der Umfang der Veränderungen im zeitlichen Muster ist gerade in den letzten Jahren jedoch so erheblich, dass eine vertiefte Beschäftigung damit angezeigt sein könnte. Insbesondere wäre es entscheidend, die Ursachen der beobachteten Verschiebung zu ermitteln (z.B. konjunkturelle Einflüsse, regulatorische Veränderungen), da nur dann eine korrekte Modellierung auch für ex-ante Prognosen erfolgen kann. Angesichts der derzeitigen Datenlage müsste dabei aber vorrangig auf theoretische Überlegungen zurückgegriffen werden; die verfügbaren Daten lassen keine weitergehende empirisch abgestützte Modellierung zu.

Zur Verbesserung der Schätzung der Steuereinnahmen scheint es uns deshalb insbesondere prüfenswert, ob die in den Verteilungsmustern enthaltenen zeitlichen Trends allenfalls besser berücksichtigt werden können, als dies bereits heute gemacht wird. Die Berücksichtigung der kantonalen Desaggregation bei der Modellierung des zeitlichen Zusammenhangs zwischen Steuererträgen und Steuereinnahmen scheint uns hingegen wenig geeignet, weil die mit ihr verbundene Verbesserung der ex-post Prognosequalität verhältnismässig geringfügig ausfiel.

Verbesserung der Datenlage

Über diese technischen Empfehlungen hinaus ist in unserer Arbeit mehrfach deutlich geworden, dass die Datenlage für die Prognose der direkten Bundessteuer auf Seiten

der Steuerdaten (Erträge und Einnahmen) spärlich ist. Es erscheint uns deshalb für zukünftige Verbesserungen der Prognosegüte sinnvoll, wenn die Datenlage systematisch verbessert würde. Dies würde unter anderem Folgendes beinhalten:

- Eine Erhebung von präzisen Daten zu den Verlustvorträgen von Unternehmen
- Nach Unternehmensmerkmalen wie Branche, Grösse und Internationalität desaggregierte Gewinnsteuerdaten (dies insbesondere deshalb, weil es sich in der Studie gezeigt hat, dass mit der Berücksichtigung des Steuerstatus, welcher als Proxy für solche Merkmale gesehen werden kann, die ex-post Prognosequalität verbessert werden kann)
- Nach natürlichen und juristischen Personen getrennte Daten zum Ablieferungsverhalten
- Grundsätzlich ist aus akademischer wie aus budgetärer Sicht anzustreben, dass alle (weiteren) nicht-sensiblen Daten, welche bisher nur den kantonalen Steuerbehörden vorliegen, schweizweit harmonisiert und so zeitnah wie möglich zugänglich gemacht werden.

5 Literatur

ESTV (2015): Schätzmethode direkte Bundessteuer, Fassung vom 1. Juli 2015, Eidgenössisches Finanzdepartement EFD, Eidgenössische Steuerverwaltung ESTV, Hauptabteilung Steuerpolitik.