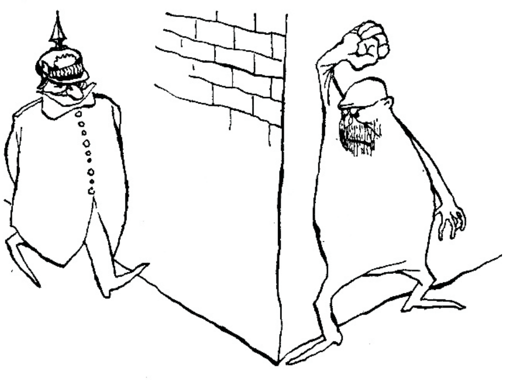
Rozdział 5. Prognozowanie – podstawowe pojęcia

Już w momencie gdy prognozujesz, wiesz że się mylisz. Nie wiesz tylko o ile i w jakim okresie.
Jedynym powodem dla którego spodziewamy się, że przyszłość będzie podobna do przeszłości jest to, że w przeszłości przyszłość była podobna do przeszłości.

[autor nieznany]

5.1. Wstęp

Prognozowanie przypomina wyobrażanie sobie tego, co jest za rogiem ulicy, jednakże bez faktycznego zaglądania za ten róg. Błędy prognoz oraz ich zróżnicowanie poglądowo zilustrował Szymon Kobyliński na rysunku, na którym każda z *dramatis personae* nietrafnie przewiduje to, co zdarzy się za rogiem:



Rysunek 5.1. Ilustracja błędów prognoz i ich zróżnicowania

Źródło: [Rolbiecki, 1970].

Prognozowanie jest rodzajem wnioskowania **ze znanego o nieznanym**. Przy danym stanie wiedzy ma ono dostarczyć jak najlepszego wyobrażenia o przyszłości. Jest to jednak zwykle wyobrażenie bardzo ułomne, z reguły znacznie mniej dokładne niż byśmy tego sobie życzyli.

Nazwą **prognozowanie** obejmujemy zarówno wnioskowanie oparte na danych czasowych, czyli prognozowanie o przyszłym biegu zdarzeń, jak też wnioskowanie oparte na danych przekrojowych, np. wnioskowanie o dochodach pewnej gminy na podstawie modelu opisującego kształtowanie dochodów innych gmin danego województwa.

Prognoza jest naukowo uzasadnionym sądem o stanie zjawiska w określonym momencie (okresie) należącym do przyszłości. Użycie w powyższym określeniu słowa „sąd" sygnalizuje niepewność prognozy, a odwołanie się do nauki oznacza, że prognoza musi być racjonalnym wnioskowaniem, prowadzącym od przesłanek do wniosków odnoszących się do przyszłości [Cieślak, 1998 s. 91].

Nie oceniamy prognoz w kategoriach **prawdziwe/fałszywe**, raczej określamy je jako **trafne**, gdy okazują się wystarczająco bliskie realizacji prognozowanej zmiennej, lub **nietrafne**, gdy rozbieżność prognozy i wielkości prognozowanej okazuje się zbyt wielka jak na nasze potrzeby.

Budowa prognozy nie jest celem samym w sobie. Tam, gdzie prognozowanie jest działaniem przygotowującym inne działanie mówimy o **preparacyjnej** funkcji prognozy. Gdy prognozy zjawisk gospodarczych przez sam fakt ich opublikowania mogą wywierać wpływ na prognozowane zjawisko, mówimy o **aktywizującej** funkcji prognozy. Prognozy aktywizujące często stają się **prognozami samorealizującymi** (np. przewidywanie inflacji wyższej niż zakładają to oficjalne dokumenty rządowe może sprawić, iż inflacja faktycznie przekroczy poziom zakładany przez rząd) bądź też **prognozami samounicestwiającymi** (np. wiosną 1995 r. prognoza ostrzegawcza o możliwości dalszego wzrostu zasobów dewizowych NBP spowodowała wprowadzenie płynnego kursu złotego w stosunku do walut zachodnich, co spowolniło dalszy wzrost tych zasobów i zmniejszyło presję inflacyjną). Wpływ prognoz na prognozowane zjawisko może znacząco obniżyć trafność prognozy. Wykorzystanie prognozy ostrzegawczej do zapobieżenia niekorzystnemu rozwojowi sytuacji może sprawić, że prognoza jest tym cenniejsza, im skuteczniej potrafi uruchomić działania zapobiegające realizacji tejże prognozy (szerzej o funkcjach prognoz w [Rolbiecki, 1970], także [Cieślak, 1998] i [Sułek, 2010] omawiają zagadnienia nieporuszane w niniejszej pracy, jak np. metoda delficka czy prognozowanie przez analogie).

Praktycznie zawsze prognozy są obarczone błędem. W prostych przypadkach prognoz budowanych w modelach statystyczno-matematycznych można sformułować przepis pozwalający na jednoczesne z prognozą wyznaczenie oczekiwanego (standardowego) błędu prognozy charakteryzującego precyzję procesu prognozowania. W miarę przechodzenia do coraz bardziej skomplikowanych modeli piętrzą się problemy związane z wyliczaniem błędów oczekiwanych. Realistyczne modele często wykazują **syndrom dużego modelu**, tzn. są współzależne, dynamiczne i nieliniowe (przynajmniej fragmentami). O ile dla wielorównaniowych, dynamicznych modeli liniowych można wyprowadzić wzory na oczekiwany błąd prognozy, o tyle dla modelu, w którym pojawia się relacja nieliniowa – nie znamy takich wzorów. Pozostaje jednak możliwość wykorzystania symulacji stochastycznej do wyznaczania oczekiwanych błędów prognozy. W dalszej części pracy przedstawiamy przegląd podstawowych problemów związanych z odpowiednim zaplanowaniem symulacji stochastycznej tak, aby nie tylko wyznaczyć oczekiwany błąd prognozy, ale i rozpoznać i skwantyfikować wpływ różnych źródeł niepewności na błąd oczekiwany.

5.2. Podstawy prognozowania

U podstaw prognozowania leży zjawisko korelacji zmiennych ekonomicznych. Jak już powiedzieliśmy, jedynym powodem, dla którego spodziewamy się, że przyszłość będzie podobna do przeszłości jest to, że *w przeszłości przyszłość była podobna do przeszłości.* Innymi słowy, fakt, że w przeszłości zaobserwowaliśmy występowanie pewnych powiązań między zmiennymi sprawia, iż – w nadziei, że związki te utrzymają się i w przyszłości – próbujemy o tej przyszłości formułować sądy, czyli prognozy.

Można wyróżnić następujące okoliczności sprawiające, iż zmienne ekonomiczne korelują ze sobą:

1. Zmienne pozostają w zależności przyczynowej – w takim przypadku staramy się opisać w modelu charakter zależności, a zwłaszcza jej kierunek, wyróżniając zmienne będące **przyczynami** oraz zmienne będące **skutkami**. Modele takie nazywamy strukturalnymi (używana jest też nazwa modele przyczynowe albo przyczynowo-skutkowe), gdyż opisują strukturę (mechanizm) powiązań pomiędzy zmiennymi. Im bardziej stabilny jest związek przyczynowy (i im dokładniej udało się nam ująć go w modelu), tym większa pewność, z jaką możemy wykorzystywać taki model do prognozowania; w szczególności w prognozowaniu możemy analizować nie tylko skutki zmian wartości zmiennych egzogenicznych, ale także skutki zmian wartości parametrów strukturalnych.

2. Zmienne pozostają w zależności symptomatycznej – w takim przypadku nie potrafimy wskazać zachodzącego między nimi związku przyczynowego, ale mamy podstawy sądzić, że istnieje ukryty mechanizm – jakaś wspólna przyczyna sprawiająca, iż zmienne niezależące od siebie zachowują się podobnie. Jeśli nie potrafimy takiej przyczyny wyrazić w postaci zmiennej liczbowej, poszukujemy zwykle zmiennej w przybliżeniu odzwierciedlającej zachowanie wspólnej przyczyny (czasem nazywamy ją zwięźle proxy), przy czym w różnych równaniach mogą wystąpić różne zmienne typu proxy. Trafność prognoz jest tu uzależniona już nie tylko od stabilności owego ukrytego mechanizmu warunkującego trwałość zależności symptomatycznej, ale i od tego, czy czynniki aktywne oraz pasywne w przeszłości pozostaną podobnie aktywne lub pasywne w przyszłości[1].

[1] W przypadku modelu przyczynowego – zwykle ekonometrycznego – dysponujemy opisem mechanizmu wiążącego zmienne, możemy więc analizować skutki uruchomienia lub zablokowania wahań poszczególnych zmiennych. W modelu symptomatycznym mechanizm wiodący od ukrytej zmiennej, będącej wspólną przyczyną, do zmiennych występujących w modelu (w szczególności do zmiennych proxy) jest ukryty.

W pewnych przypadkach nie potrafimy wskazać ani powodów istnienia ukrytego mechanizmu, ani też odpowiedniej zmiennej proxy, ale:

1. Jeśli rozważane zmienne charakteryzują się skłonnością do jednokierunkowych zmian w czasie, wówczas mówimy, że wspólną „przyczyną” takiej skłonności jest trend czasowy, konstruujemy odpowiednią zmienną sztuczną, możliwie prostą, ale pozwalającą opisać tendencję zmian zmiennej objaśnianej. Walory poznawcze **modeli trendu** (zwanych też modelami tendencji rozwojowej) są ograniczone tym, że nazywając nieznaną przyczynę trendem, niewiele posuwamy się naprzód w wyjaśnieniu jej natury. Modele trendu są jednak popularne ze względu na łatwość ich budowy i wykorzystania do prognozowania.
2. Jeśli modelowana zmienna wykazuje wyraźne wahania, wówczas możemy założyć, że informacja o działaniu ukrytego mechanizmu jest zawarta w zrealizowanych wartościach modelowanych zmiennych. Budujemy wówczas różne modele autoregresyjne (oznaczane, stosownie do stopnia ich komplikacji, symbolami AR, ARMA, ARIMA) lub szerzej – modele szeregów czasowych.
3. Zmienne korelują ze sobą przypadkowo – nawet jeśli korelacja taka jest silna, stosowanie modeli zawierających przypadkowo skorelowane zmienne jest nadzwyczaj ryzykowne, a właściwe bezzasadne. Mimo to niejednokrotnie próbuje się wykorzystać korelację przypadkową do prognozowania. Kryje się za tym cicha nadzieja, że ukryty mechanizm jednak istnieje; działa tu fakt, że *jedynym powodem, dla którego spodziewamy się, że przyszłość będzie podobna do przeszłości jest to, że w przeszłości przyszłość była podobna do przeszłości*.

Czasem stwierdzenie istnienia korelacji przypadkowej staje się bodźcem do owocnego poszukiwania ukrytego mechanizmu, odpowiedzialnego za tę korelację, **częściej** jednak prowadzi do powstania niedobrego modelu.

5.2.1. Prognozowanie strukturalne i niestrukturalne

W podejściu **strukturalnym** instrumentem prognozowania jest model strukturalny, tj. model, którego równania zostały skonstruowane tak, aby odzwierciedlić pewną teorię, stanowić opis mechanizmu generującego prognozowane wielkości. Estymacja parametrów modelu strukturalnego dostarcza ocen parametrów oraz pozwala na weryfikację zgodności teorii, która legła u podstaw modelu, z danymi statystycznymi. Właściwe wykorzystanie teorii ekonomicznej wnosi dodatkową informację, informację **a priori** w stosunku do informacji zawartej w próbie statystycznej, uzupełniając w ten sposób nie zawsze obfitą informację statystyczną. Modelom strukturalnym stawia się zarzut wykorzystywania nadmiernie uproszczonej teorii ekonomicznej, co jest często odzwierciedleniem lakonicznego, mało precyzyjnego charakteru samej teorii [Sims, 1980], [Ta Chung Liu, 1960].

W podejściu **niestrukturalnym** przyjmuje się, że w przeszłości prognozowanej zmiennej jest zawarta informacja o mechanizmie, który zmienną kształtuje. W metodach **naiwnych** oraz metodach **wyrównywania wykładniczego** prognozowanie ma charakter dość mechaniczny, co pozwala na wykorzystanie stosunkowo mało licznej próby. W metodach **analizy szeregów czasowych** staramy się uchwycić dość subtelny, wewnętrzny mechanizm powodujący, iż przeszłość szeregu czasowego wpływa na jego przyszłość. Estymacja parametrów modelu szeregu czasowego dostarcza ocen parametrów i informuje przy tym, jak daleko wstecz sięga mechanizm uzależniający teraźniejszość od przeszłości. Głównym źródłem informacji jest tu próba statystyczna, która winna być odpowiednio liczna, aby zapewnić wystarczającą dokładność estymatorów. Za niedoskonałości modeli analizy szeregów czasowych uważa się wysokie wymagania co do liczebności próby. Modelom niestrukturalnym wytyka się też abstrahowanie od ważnego źródła informacji, jakim jest teoria ekonomiczna będąca skumulowaną w czasie informacją o modelowanym systemie.

W praktyce wykorzystujemy mieszankę obydwu podejść. Budując model strukturalny, wielokrotnie dochodzimy do wniosku, że teoria ekonomiczna jest zbyt mało precyzyjna i nie pozwala na jednoznaczne wyspecyfikowanie równania przeznaczonego do oszacowania. Kierując się zasadami teorii ekonomicznej, ostateczną postać równania określamy na podstawie analizy korelacji pomiędzy realizacjami zmiennych, jakie trafiły do próby statystycznej. Postępujemy w sposób zbliżony do podejścia niestrukturalnego. Z drugiej strony – w modelach niestrukturalnych, zwłaszcza zaawansowanych wielorównaniowych modelach typu VAR, szereg decyzji dotyczących np. doboru zmiennych czy też zaklasyfikowania ich do klasy zmiennych endogenicznych lub egzogenicznych jest motywowane aprioryczną wiedzą o modelowanym systemie. Ponadto w obydwu typach modeli często występuje zmienna czasowa reprezentująca tendencję do systematycznych, jednokierunkowych zmian. W sumie podejścia te można uznać raczej za komplementarne niż konkurencyjne, o wyborze zaś jednego z nich decydują takie okoliczności, jak zasób dostępnej informacji a priori przyjmującej postać teorii ekonomicznej oraz zasób dostępnej informacji statystycznej, zawartej w próbie.

5.2.2. Etapy prognozowania

Proces prognozowania można podzielić na **etapy**. Przedstawimy ich klasyfikację wzorowaną na pracy [Cieślak, 1998]:

1. Sformułowanie zadania prognostycznego – określamy tu:
   - obiekt,
   - zjawisko,
   - prognozowane zmienne,
   - cel wyznaczenia prognozy, pożądany horyzont prognozy
   - wymagania co do dokładności prognozy.
2. Sformułowanie przesłanek prognozy – określamy tu:
   - opis mechanizmu generującego prognozowaną zmienną,

 - listę czynników wpływających na ten mechanizm (w szczególności w modelu ekonometrycznym zaliczymy tu listę zmiennych egzogenicznych wpływających na modelowane zjawiska).

3. Wybór predyktora – przepisu, według którego wyznaczamy prognozę oraz zasady prognozowania. Na wybór ten składają się decyzje co do metody prognozowania (na podstawie modelu formalnego strukturalnego lub niestrukturalnego, opinii ekspertów, metodą heurystyczną, analogową itp.).
4. Wyznaczenie prognozy – realizacja tego etapu jest związana bezpośrednio z wyborem sposobu i zasady prognozowania. Zastosowanie modelu formalnego oznacza konieczność zbudowania nowego modelu lub wykorzystania (i/lub adaptacji) modelu istniejącego; prognozowanie oparte na opinii ekspertów wymaga określenia grona ekspertów, zapewnienia sobie ich współpracy, określenia sposobu przekształcenia opinii ekspertów (w razie gdyby były zróżnicowana lub wręcz rozbieżne) w jednoznaczną prognozę itd.
5. Ocena dokładności prognozy – rozróżniamy tu ocenę ex post opartą na zrealizowanym błędzie prognozy, dokonywaną po zrealizowaniu się zmiennej prognozowanej (tj. gdy prognoza wygaśnie) oraz ocenę ex ante opartą na oczekiwanym błędzie prognozy – wyrażającą przypuszczenia co do dokładności prognozy wyrażone zanim prognoza wygaśnie.

5.2.3. Warunki prognozowania z modelu ekonometrycznego

Analiza zasadności zastosowania modelu do prognozowania sprowadza się do udzielenia odpowiedzi na następujące pytania:

- Czy model jest stabilny w okresie prognozy? Rozważamy tu możliwość zaistnienia okoliczności mogących spowodować dezaktualizację modelu, np. wskutek zmiany mechanizmu generującego wartości zmiennych endogenicznych czy to przez zmianę jego postaci funkcyjnej, czy też przez zmianę zestawu zmiennych objaśniających.
- Czy parametry strukturalne modelu są stabilne w okresie prognozy? Rozważamy tu możliwość zaistnienia okoliczności mogących spowodować zmianę parametrów mechanizmu generującego wartości zmiennych endogenicznych. Przykładem sytuacji budzącej obawy, iż doszło do zmiany parametrów, może być wpływ zmiany stopy podatkowej na parametry równania opisującego udział podatków w dochodzie narodowym. Zmiana parametrów modelu jest określana zazwyczaj mianem zmiany strukturalnej. Do najprostszych metod uwzględniania zmian strukturalnych zaliczyć można stosowanie zmiennych zerojedynkowych zmieniających w skokowy sposób wartości parametrów modelu. Istnieje też obszerna klasa modeli o zmiennych parametrach (por. [Dziechciarz, 1994]).
- Czy parametry struktury stochastycznej modelu (parametry rozkładu zakłóceń) są stabilne? Rozważamy tu możliwość zaistnienia okoliczności mogących

spowodować dezaktualizację opisu sposobu, w jaki zakłócenia modelu wpływają na wartości zmiennej endogenicznej.

Przykładem sytuacji budzącej obawy, iż doszło do zmiany mechanizmu może być zmiana podziału administracyjnego kraju w kontekście równania opisującego stopę bezrobocia. W równaniu opisującym poziom bezrobocia w nowym, powiększonym województwie, zakłócenia zapewne będą miały większą wariancję (choć w równaniu opisującym stopę bezrobocia, a więc zmienną wyrażającą intensywność zjawiska, zakłócenia mogą mieć rozkład o niezmienionych parametrach.

Zwykle znacznie łatwiej jest przedstawić argumenty na to, że zaszły zmiany strukturalne niż na rzecz stabilności modelu, jego parametrów tudzież zakłóceń.

Posiłkujemy się zatem praktyczną zasadą braku wystarczającego powodu sformułowaną przez Laplace'a: jeśli z analizy modelu nie wynikają wystarczająco uzasadnione obawy o to, że model, jego parametry lub zakłócenia są niestabilne – przyjmujemy, iż są one stabilne.

5.2.4. Zasady prognozowania[2]

Przed przystąpieniem do prognozowania (zwłaszcza przy wykorzystaniu modelu ekonometrycznego) musimy rozstrzygnąć, według jakiej zasady ma być skonstruowany przepis zwany predyktorem, który – po podstawieniu do niego wartości zmiennych objaśniających, wygeneruje prognozę. Wybór tej czy innej zasady prognozowania jest uzależniony od tego, ile „płacimy" za błędy prognoz, tj. jak wygląda nasza funkcja strat. Do głównych zasad prognozowania należą:

1. Zasada prognozowania według wartości oczekiwanej – jest to najpopularniejsza z zasad prognozowania. Jej popularność wiąże się z dwoma zaletami – statystyczną: daje prognozy nieobciążone, oraz praktyczną: jest łatwa do wyznaczenia. Predyktorem jest warunkowa wartość oczekiwana, tj. regresja zmiennej objaśnianej względem zmiennych objaśniających

$$\hat{y}_t = E(y_t | x_{1t}, x_{2t} \ldots), \quad t = 1, \ldots, T. \tag{5.1}$$

   Pojedyncze prognozy otrzymane według tej zasady będą, jak każda prognoza, na ogół różniły się od przyszłej realizacji prognozowanej zmiennej. Jeśli jednak będziemy wielokrotnie powtarzać prognozowanie przebiegu zjawiska w tych samych warunkach, należy oczekiwać, że średni błąd ciągu tak otrzymanych prognoz będzie równy 0. Zasada ta daje prognozy minimalizujące **błąd średniokwadratowy**:

$$S = 1/T \sum_{t=1}^{T} (y_t - \hat{y}_t)^2. \tag{5.2}$$

[2] W literaturze używa się też terminu zasady predykcji na określenie zasady budowy prognozy [Pawłowski, 1975], [Zeliaś, 1993]; w pracy [Czerwiński, 1980] autor używa nazwy reguły prognozowania. Aby uniknąć stosowania nadmiaru terminów, unikamy używania słowa predykcja, choć w języku angielskim funkcjonuje nawet rozróżnienie pojęć *forecasting* (prognoza w przód) oraz *backcasting* (prognoza wstecz).

2. Zasada prognozowania według **największego prawdopodobieństwa** – w naukach nieeksperymentalnych (a taką jest ekonomia) trudno znaleźć sytuację, w której można wielokrotnie prognozować zjawiska zachodzące w tych samych warunkach. Dużo częściej prognoza ekonomiczna ma charakter incydentalny, dla danych warunków jest wyznaczana tylko raz. W takim przypadku bardziej użyteczna może okazać się prognoza mająca najwyższe prawdopodobieństwo realizacji, tj. prognoza na poziomie **dominanty** (**mody**) rozkładu prognoz
$$\hat{y}_t = \left\{ \hat{y}_t^D \colon \Pr\left(\hat{y}_t^D\right) \geqslant \Pr\left(\hat{y}_t^o\right) \right\} \quad t = 1, \ldots, T, \tag{5.3}$$
gdzie $\Pr\{\hat{y}_t^o\}$ jest prawdopodobieństwem (w przypadku zmiennych ciągłych $\Pr\{\hat{y}_t^o\}$ jest wartością funkcji gęstości w punkcie $\{\hat{y}_t^o\}$), z jakim zmienna prognozowana przyjmie wartość y_t^o , różną od dominanty y_t^D zmiennej y_t. Zasada ta daje prognozy maksymalizujące prawdopodobieństwo trafienia w przyszłą realizację zmiennej, a funkcja strat
$$S = -\max_{\varepsilon \to 0^+} \Pr\left\{ \left| \hat{y}_t - y_t^D \right| \right\} < \varepsilon \tag{5.4}$$
osiąga minimum, gdy prognoza $\hat{y}$ przyjmie wartość dominanty y^D rozkładu. Prognoza najbardziej prawdopodobna jest bardzo atrakcyjną alternatywą dla prognozy nieobciążonej. W przypadku rozkładu wyraźnie asymetrycznego może jednak dać skrajne wartości zmiennej. W praktyce modelowania ekonometrycznego stosuje się ją rzadko ze względu na trudności obliczeniowe. Bardzo pomocna może się tu okazać symulacja stochastyczna.
3. Zasada prognozowania według **mediany** – wykorzystuje trzecią (obok wartości oczekiwanej i dominanty) charakterystykę centrum rozkładu prognozowanej zmiennej, czyli medianę
$$\hat{y}_t = \left\{ \hat{y}_t^M \colon \Pr\left(\hat{y}_t^M \leqslant \hat{y}_t^o\right) = 0,5 \wedge \Pr\left(\hat{y}_t^M \geqslant \hat{y}_t^o\right) = 0,5 \right\}, \quad t = 1, \ldots, T, \tag{5.5}$$
gdzie y_t^o jest wartością zmiennej prognozowanej różną od mediany y_t^M zmiennej y_t.
Zasada ta jest stosowana np. wówczas, gdy w charakterze prognozy wykorzystujemy **rozwiązanie deterministyczne** modelu wielorównaniowego o szczególnej budowie, zapewniającej wzajemnie jednoznaczne odwzorowanie zakłóceń u w zmienne endogeniczne y (por. [Hall, 1985b]). Również tutaj pomocna może się okazać symulacja stochastyczna. Zasada ta, szczególnie użyteczna w przypadku rozkładów z grubymi ogonami, daje prognozy minimalizujące **średni błąd bezwzględny**:
$$S = 1/T \sum_{t=1}^{T} \left| y_t - \hat{y}_t \right|. \tag{5.6}$$
4. Zasada prognozowania **minimalizującego oczekiwaną stratę** – jest stosowana w sytuacjach, gdy błędowi prognozy $(y - \hat{y}_t^s)$ możemy przypisać określoną

stratę $W(\hat{y}_t^s)$; wybieramy wówczas prognozę związaną z najmniejszą oczekiwaną stratą

$$\hat{y}_t = \{\hat{y}_t^s\colon W(\hat{y}_t^s) \leqslant W(\hat{y}_t^o)\} \tag{5.7}$$

Przykładem prognozy skonstruowanej według tej zasady może być prognoza wyznaczona w wyniku optymalnego sterowania modelem ekonometrycznym (por. [Gajda, 1993]).

5.3. Prognozy ex post i ex ante oraz ich błędy

Prognoza ekonometryczna jest sądem warunkowym typu: *jeżeli iks przybierze wartość x^o, to spodziewamy się, że igrek przybierze wartość y^o*. Warunkowy charakter prognozy uwidocznia się szczególnie wyraźnie w tzw. **prognozach kontrfaktycznych**, w których rozważając przeszłość zadajemy sobie pytanie: jaki byłby przebieg wydarzeń, gdyby zmienne egzogeniczne przyjęły wartości odmienne od wielkości faktycznie zrealizowanych. Prognozy wyznaczane z modelu ekonometrycznego różnicujemy ze względu na zasób wykorzystanej informacji o zmiennych egzogenicznych na:

a) **prognozy ex post** budowane na podstawie znanych wartości zmiennych egzogenicznych,
b) **prognozy ex ante** budowane na podstawie nieznanych wartości zmiennych egzogenicznych – wartości, które trzeba zaprognozować lub wyznaczyć w inny sposób, aby następnie wyznaczyć właściwą prognozę.

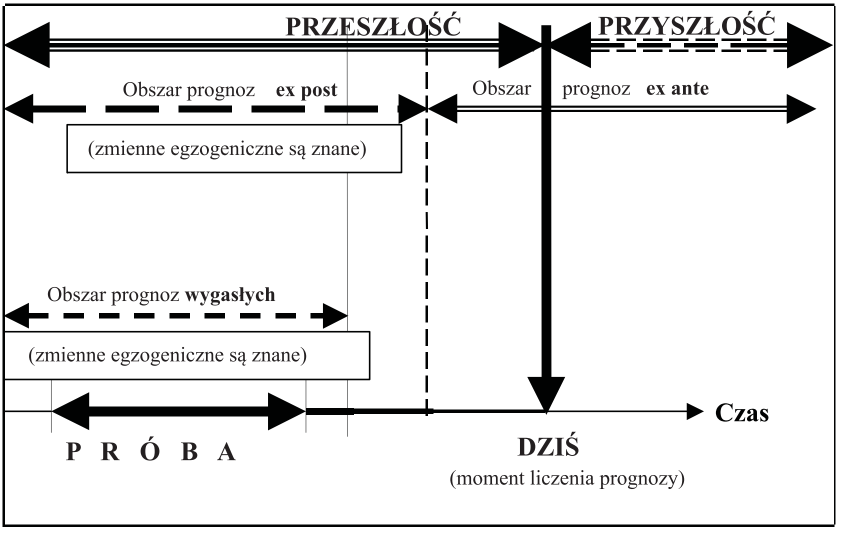
Autorzy pracy [Intriligator, Bodkin, Hsiao, 1996, s. 520] podkreślają, że prognoza ekonometryczna bazująca na prognozowanych zmiennych egzogenicznych $\hat{\mathbf{z}}_{T+1}$ jest nazywana ex ante, ponieważ jest prawdziwym przewidywaniem (*true forecast*), **zanim** pojawi się zdarzenie. W odróżnieniu od niej prognoza ex post, realizowana po zdarzeniu, zastępuje przewidywane wartości zmiennych egzogenicznych $\hat{\mathbf{z}}_{T+1}$ ich wartościami zrealizowanymi $\mathbf{z}_{T+1}$, a poprawki stałych $\hat{\mathbf{u}}_{T+1}$ zerowymi wartościami oczekiwanymi zakłóceń. Wypada nadmienić, że niektórzy autorzy (por. [Welfe, 1995, s.180]) nazywają prognozę ex ante prognozą warunkową, zaś prognozę ex post prognozą bezwarunkową.

Wartości zmiennych egzogenicznych modelu ekonometrycznego musimy wcześniej ustalić poza modelem. W praktyce prognozy ex post wyznaczamy dla prognoz wygasłych, dla których znamy realizacje zmiennych egzogenicznych. Aby wyznaczyć prognozy ex ante, trzeba najpierw wyznaczyć spodziewane wartości zmiennych egzogenicznych, korzystając z obcych badań, niestrukturalnych metod prognozowania lub też kierując się wyczuciem badacza, który w takim momencie staje się ekspertem sam dla siebie.

Błąd prognozy ex post zależy głównie od własności modelu, podczas gdy na błąd prognozy ex ante wpływają zarówno własności modelu, jak i trafność założeń co do wartości zmiennych egzogenicznych.

Ze względu na moment wyznaczenia błędu prognozy odróżniamy błąd zrealizowany prognozy od oczekiwanego błędu prognozy. Po zrealizowaniu wartości zmiennej prognozowanej prognoza wygasa. W przypadku prognoz wygasłych możemy wyznaczyć zrealizowany błąd prognozy $\hat{u}_t = y_t - \hat{y}_t$ (wyliczane w procesie estymacji reszty są najczęściej spotykanym przykładem błędów prognoz wygasłych). Zaletą tej miary jest to, że **zrealizowany błąd prognozy** podaje kompletną informację o dokładności prognozy. Istotną wadą jest to, że informacja ta staje się dostępna dopiero wówczas, gdy prognoza **wygasa**. Wady tej nie ma **oczekiwany błąd prognozy**. W celu odróżnienia od błędu zrealizowanego – błąd oczekiwany jest nazywany czasem **błędem predykcji** (por. [Pawłowski, 1975], [Zeliaś 2000], [Czerwiński, Guzik, 1980]). O użyteczności oczekiwanego błędu prognozy decyduje fakt, że może on być wyliczony jednocześnie z prognozą, pozwala więc ocenić dokładność prognozy ex ante w momencie budowy prognozy, w którym prognoza może służyć podejmowaniu decyzji. Jego wadą jest zaś to, że jako charakterystyka rozproszenia błąd prognozy mówi, o ile plus/minus możemy się pomylić, nie mówi jednak o kierunku błędu i jego absolutnej wartości.

Relacje pomiędzy omówionymi powyżej pojęciami zilustrowano na rysunku.



Rysunek 5.2. Relacje pomiędzy rodzajami prognoz

Źródło: [Gajda, 1988].

Rozdział 6. Prognozowanie z modeli strukturalnych

6.1. Prognozowanie z modeli jednorównaniowych

Oszacowania parametrów równań modeli wykorzystywanych do prognozowania zazwyczaj są znane z wcześniejszych etapów budowy modelu. Podejmując decyzję o wykorzystaniu modelu do prognozowania, możemy jednak rozważyć, czy wykorzystana metoda estymacji daje oceny najlepsze z punktu widzenia dokładności prognozy. W szczególności wypada tu wspomnieć o dokonanym przez H. Wolda rozróżnieniu metod estymacji na zorientowane na precyzję prognoz (np. metoda najmniejszych kwadratów) oraz metody zorientowane na precyzję ocen parametrów strukturalnych (np. metoda największej wiarygodności). Te dwa kryteria nie muszą prowadzić do estymatorów o tych samych własnościach.

Szczególnie często stosuje się estymator metody najmniejszych kwadratów. Prognozowanie z modeli oszacowanych za pomocą metody najmniejszych kwadratów, acz w modelach współzależnych metodologicznie niezbyt zasadne, w praktyce ekonometrycznej jest wręcz nagminne. Częstym uzasadnieniem jest tu argument o wyższej precyzji prognoz opartych na ocenach metody najmniejszych kwadratów. Stąd podstawowe wzory cytujemy w wersji dla estymatora MNK.

6.1.1. Oczekiwany błąd prognozy w modelu z jedną zmienną objaśniającą

Rozważmy oczekiwany błąd prognozy ex post z modelu jednorównaniowego z jedną zmienną objaśniającą

$$y_t = b_1 + b_2 x_t + u_t, \quad t = 1, \dots, T \tag{6.1}$$

wyznaczany zgodnie z formułą

$$\hat{y}_\tau = \hat{b}_1 + \hat{b}_2 x_\tau + 0, \quad \tau > T, \tag{6.2}$$

gdzie wartość zmiennej objaśniającej to x_τ, a zakłócenia u_τ przyjmujemy na poziomie zerowym (ich wartości oczekiwanej).

Czasem jednak w zrealizowanych resztach kryje się użyteczna informacja pozwalająca na podniesienie precyzji prognozy. Może to być np. ujawniona autokorelacja reszt, dzięki której można wprowadzić do formuły predyktor różnej od zera wartości przewidywanego zakłócenia $\hat{u}_t$.

W szczególności pojawienie się w okresach poprzedzających okres prognozy serii reszt o jednakowych znakach może być interpretowane jako sygnał tego, że już po oszacowaniu parametrów równania w jego strukturze zaszła zmiana. Z braku lepszego argumentu przyjmuje się wtedy, że zmiana ta spowodowała przesuniecie całego równania w górę lub w dół i modyfikuje wyraz wolny o pewną wielkość, zwaną *constant adjustment*. Są to tzw. **poprawki stałych** (tłumaczenie alternatywne to stałe poprawki, choć w wielu przypadkach poprawki te w okresie prognozy wcale nie są stałe). Kwestia stosowania tego typu modyfikacji pozostaje otwarta, *constant adjustments* potrafią bowiem dramatycznie zdeformować relacje pomiędzy zmiennymi modelu, niszcząc tym samym racjonalność modelu. Może to być przyczyną tego, że autorzy prognoz nie mają zwyczaju publikowania razem z prognozami *constant adjustments* przyjętych przy wyznaczaniu prognoz.

Praktyka pokazuje, ze można tak dobrać te poprawki, aby prognozy „generowane przez model" biegły dokładnie po trajektorii oczekiwanej przez autora prognoz, odzwierciedlając oczekiwania tegoż producenta, a jednocześnie nie miały nic wspólnego z samym modelem, por. [Klein, 1971, 1982], [Gajda 1988, 2004], a także pracę [Fair, 1994], gdzie autor zajmuje szczególnie krytyczne stanowisko i optuje za reestymacją równań w miarę napływu świeżych danych statystycznych, por. też [Clements, Hendry 1998a].

Błąd prognozy ex post wynosi

$$\begin{aligned} y_\tau - \hat{y}_\tau &= (b_1 + b_2 x_\tau + u_\tau) - (\hat{b}_1 + \hat{b}_2 x_\tau) \\ &= (b_1 - \hat{b}_1) + (b_2 - \hat{b}_2)x_\tau + u_\tau, \qquad \tau > T \end{aligned} \tag{6.3}$$

zależy zatem od błędów estymacji parametrów strukturalnych $b_i - \hat{b}_i$ oraz od zakłóceń u_τ równania wykorzystanego do prognozowania.

Jeżeli parametry oszacowano estymatorem MNK przy klasycznych założeniach – przypomnijmy, że zakładamy:

– zerową wartość oczekiwaną zakłóceń $\mathrm{E}(u_t) = 0$,
– stałą wariancję zakłóceń $\mathrm{var}(u_t) = \sigma^2$,
– zerową kowariancję zakłóceń z różnych okresów $\mathrm{cov}(u_t, u_{t+\tau})$ tj. brak autokoralacji (w modelach opartych na szeregach czasowych dla prób przekrojowych założenie to nie ma dobrze określonego sensu),
– nielosowość zmiennych objaśniających,

to wówczas **oczekiwany błąd prognozy** S_τ wyznaczony jako pierwiastek z wariancji błędu prognozy ex post na okres τ wyliczamy z formuły

$$S_\tau = \sqrt{\mathrm{E}(y_\tau - \hat{y}_\tau)^2} = \sqrt{\sigma_u^2 \left[1 + \frac{1}{T} + \frac{(x_\tau - \bar{x})^2}{\sum\limits_{t=1}^{T}(x_t - \bar{x})^2}\right]}. \tag{6.4}$$

Zatem oczekiwany błąd prognozy ex post:

1) rośnie, gdy wzrasta wariancja zakłóceń σ^2,

2) rośnie, gdy wartość zmiennej objaśniającej w okresie prognozy x_τ oddala się od średniej wartości tej zmiennej $\bar{x}$ w próbie,
3) maleje, gdy wzrasta zmienność $\sum\limits_{t=1}^{T}(x_i - \bar{x})^2$ (a więc i wariancja) zmiennej objaśniającej w próbie,
4) maleje, gdy wzrasta T – liczba obserwacji w próbie.

Rozważmy teraz oczekiwany błąd prognozy ex ante. Prognozę ex ante z modelu jednorównaniowego z jedną zmienną objaśniającą (6.1) wyznaczamy zgodnie z formułą

$$\hat{y}_\tau = \hat{b}_1 + \hat{b}_2\hat{x}_\tau, \quad \tau > T, \tag{6.5}$$

gdzie w miejsce wartości zmiennej objaśnianej wstawiliśmy jej **prognozę**

$$\hat{x}_\tau = x_\tau + v_\tau, \quad \tau > T. \tag{6.6}$$

Dla uproszczenia rozważań załóżmy, że błąd prognozy v_τ zmiennej objaśniającej x_τ ma zerową wartość średnią i stałą wariancję $\sigma^2_{v_\tau}$, a ponadto jest nieskorelowany z zakłóceniem u_τ, dzięki czemu prognoza pozostanie nieobciążona.

Błąd prognozy ex ante wynosi

$$\begin{aligned} y_t - \hat{y}_t &= (b_1 + b_2x_\tau + u_\tau) - (\hat{b}_1 + \hat{b}_2\hat{x}_\tau) \\ &= (b_1 - \hat{b}_1) + (b_2 - \hat{b}_2)x_\tau + (u_\tau - \hat{b}_2v_\tau), \quad \tau > T \end{aligned} \tag{6.7}$$

jest więc równy błędowi prognozy ex post powiększonemu o wyrażenie $(u_\tau - \hat{b}_2v_\tau)$. Błąd prognozy ex ante może zatem być większy lub mniejszy od błędu prognozy ex post w zależności od znaków zakłóceń u i v oraz oceny parametru b_2. Jeżeli parametry równania oszacowano estymatorem MNK, wzór na oczekiwany błąd prognozy S_τ wyznaczony jako **pierwiastek z wariancji błędu** prognozy na okres τ przyjmie postać

$$S_\tau = \sqrt{\mathrm{E}(y_\tau - \hat{y}_\tau)^2} = \sqrt{\sigma^2_u\left[1 + \frac{1}{T} + \frac{(x_\tau - \bar{x})^2 + \sigma^2_{v_\tau}}{\sum\limits_{t=1}^{T}(x_t - \bar{x}_t)^2}\right] + \hat{b}^2_2\sigma^2_{v_\tau}} \tag{6.8}$$

Zatem oczekiwany błąd prognozy ex post:
1) rośnie, gdy wzrasta wariancja zakłóceń σ^2,
2) rośnie, gdy wartość zmiennej objaśniającej w okresie prognozy x_τ oddala się od średniej wartości tej zmiennej $\bar{x}$ w próbie,
3) maleje, gdy wzrasta zmienność $\sum\limits_{i=1}^{T}(x_i - \bar{x})^2$ (a więc i wariancja) zmiennej objaśniającej w próbie,
4) maleje, gdy wzrasta T – liczba obserwacji w próbie,
5) rośnie, gdy rośnie wariancja $\sigma^2_{v_\tau}$ błędu (v_τ), z jakim przewidujemy zmienną objaśniajacą,

6) rośnie, gdy rośnie kwadrat oceny parametru stojącego przy zmiennej egzogenicznej b_2.

Zatem wariancja oczekiwanego błędu prognozy ex ante zależy dodatkowo od wariancji błędu v, z jakim przewidujemy wartość zmiennej objaśniającej x_T, oraz od oceny parametru b_2 stojącego przy tej zmiennej.

Pierwiastek z wariancji, czyli oczekiwany błąd prognozy, mógłby być mniejszy dla prognozy ex ante niż dla prognozy ex post, jeśli błąd u prognozy igreka oraz błąd v prognozy iksa byłyby odpowiednio skorelowane i wzajemnie kompensowały swoje wpływy: musiałyby korelować ujemnie, gdyby ocena parametru b_2 była ujemna, bądź korelować dodatnio w przypadku oceny dodatniej. Uzyskanie tak szczególnej korelacji jest jednakże trudne, toteż oczekiwany błąd prognozy ex ante jest zwykle większy niż oczekiwany błąd prognozy ex post. W symulacjach na modelach serii W prowadzonych w latach 1978–1981 ok 30% zmiennych charakteryzowało się mniejszym błędem prognozy ex ante niż prognozy ex post.

6.1.2. Przedział ufności prognozy

Wypowiedź o oczekiwanej wartości zmiennej można uściślić, formułując **przedział ufności** – przedział, do którego ze znanym prawdopodobieństwem wpada zrealizowana wartość badanej zmiennej. Przydatna jest wówczas znajomość rozproszenia rozkładu zmiennej mierzona odchyleniem standardowym. W szczególności zatem znajomość **oczekiwanego kwadratu błędu prognozy** może być wykorzystana do budowy przedziału ufności dla prognozy (nazywanego czasem prognozą przedziałową) – przedziału, który ze znanym prawdopodobieństwem **pokryje** wartość prognozowanej zmiennej.

W ogólnym przypadku użyteczna okazuje się nierówności Czebyszewa

$$P\{\bar{x} - kS_s < x < \bar{x} + kS_x\} \geqslant \left(1 - \frac{1}{k^2}\right). \tag{6.9}$$

W przypadku zmiennej x ze średnią $\bar{x}$ i odchyleniem standardowym S_x oraz dla $k \leqslant 1$ nierówność ta nie daje użytecznej informacji, natomiast dla $k > 1$ pozwala ona na wyznaczenie dla zmiennej x o dowolnym rozkładzie przedziału ufności $\{\bar{x} - kS_x, \bar{x} + kS_x\}$, który nakrywa wartość zmiennej x z prawdopodobieństwem (ufnością) nie mniejszym od $\left(1 - \frac{1}{k^2}\right)$.

Jeśli zakłócenia predyktora mają rozkład normalny i mamy podstawy, aby sądzić, iż rozkład ten utrzyma się w okresie prognozy – wówczas błędy prognozy mają rozkład t-Studenta, zaś przedział ufności prognozy przyjmuje postać

$$P\{\bar{x} - t_{\alpha/2}S_x \leqslant x \leqslant \bar{x} + t_{\alpha/2}S_x\} = 1 - \alpha, \tag{6.10}$$

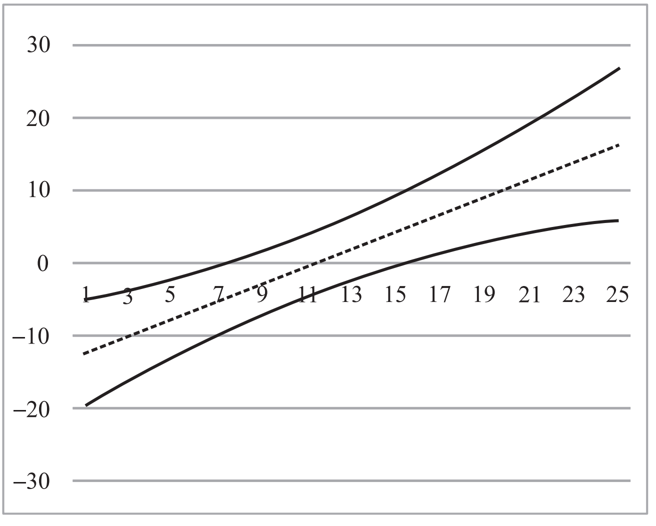
gdzie współczynnik t odczytujemy z tablic rozkładu t-Studenta.

Dla przykładu pokażemy poziomy ufności przedziałów opartych na nierówności Czebyszewa, na rozkładach t-Studenta z 6 i 60 stopniami swobody oraz na rozkładzie normalnym.

Tabela 6.1. Prawdopodobieństwo, z jakim przedział $\{\bar{x} - kS_x < x < \bar{x} + kS_x\}$ pokrywa wartość zmiennej x

Prawdopodobieństwo z jakim przedział k-sigmowy nakrywa zmienną:	**dla $k = 1$**	**dla $k = 1{,}5$**	**dla $k = 2$**	**dla $k = 2{,}5$**	**dla $k = 3$**
wg Czebyszewa co najmniej	0,000	0,55586	0,750	0,840	0,8889
dla rozkładu t-Studenta z 6 stopniami swobody	0,6441	0,8157	0,9076	0,9535	0,9760
dla rozkładu t-Studenta z 60 stopniami swobody	0,6787	0,8611	0,9500	0,9848	0,9961
dla rozkładu normalnego	0,6827	0,8664	0,9545	0,9876	0,9973

Źródło: Opracowanie własne – wyliczono w programie Excel, korzystając z funkcji statystycznych `ROZKLAD.NORMALNY.S` oraz `ROZKLAD.T`.



Rysunek 6.1. Linia regresji oraz przedział ufności prognozy wyznaczony w odległości $\pm 2\sigma_u$

Źródło: opracowanie własne.

W naszym przykładzie na wykresie pokazujemy tzw. **dwusigmowy** przedział ufności, tzn. przedział, którego krańce są oddalone o $\pm 2 \cdot \sigma$, gdzie sigma jest oczekiwanym błędem prognozy. Zgodnie z nierównością Czebyszewa, przedział taki pokryje wartość prognozowanej zmiennej z prawdopodobieństwem co najmniej $1 - \frac{1}{2^2} = 0{,}75$. W przypadku, gdy możemy skorzystać z rozkładu t-Studenta, zamiast mnożnika 2 bierzemy wartość krytyczną z tablic (powiedzmy $t_{0{,}05/2,df}$ dla poziomu istotności $\alpha = 0{,}05$ oraz liczbie stopni swobody df) i otrzymujemy przedział ufności o krańcach oddalonych o $\pm t_{0{,}05/2,df} \cdot \sigma$, który pokryje prognozowaną zmienną z ufnością 0,95. Z powyższej tabeli widać, jak cenna jest sytuacja, w której rozkład błędów prognozy jest znany. W przypadku, gdy jest to rozkład normalny lub t-Studenta, dwusigmowy przedział ufności obejmie od 90% do 95% realizacji prognozowanej zmiennej. Jeśli natomiast nie mamy informacji o rozkładzie błędów prognozy, zadowalamy się wnioskiem wypływającym z nierówności Czebyszewa, że ten sam dwusigmowy przedział pokryje **co najmniej** 75% realizacji prognozowanej zmiennej. Omawiana poniżej symulacja stochastyczna przychodzi nam tu z pomocą, pozwalając na wyznaczenie empirycznych przedziałów ufności zbudowanych np. na kwantylach wysymulowanego rozkładu błędów. Pomoc okazana przez symulację stochastyczną może okazać się szczególnie cenna w przypadku, gdy rozkład błędów jest asymetryczny.

6.1.3. Oczekiwany błąd prognozy w modelu z wieloma zmiennymi objaśniającymi

Rozważmy problem prognozowania z jednorównaniowego modelu liniowego z wieloma zmiennymi objaśniającymi

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{u}. \tag{6.11}$$

Prognozę wyliczymy zgodnie z formułą

$$\hat{y}_\tau = \hat{\mathbf{x}}_\tau \hat{\mathbf{b}} + \hat{u}_\tau, \quad \tau > T, \tag{6.12}$$

gdzie symbol $\mathbf{x}_\tau$ oznacza wiersz τ macierzy $\mathbf{X}$. W przypadku prognozy ex ante również nad symbolem zmiennej objaśniającej wstawiliśmy daszek przypominający, iż wartości zmiennych objaśniających są niedokładne, obarczone pewnym błędem: zostały zaprognozowane, wzięte z innych źródeł lub ustalone na mocy założenia.

Z zakłóceniem $\hat{u}_t$, $t > T$, sprawa jest stosunkowo prosta: w prognozowaniu często przyjmuje się wartość zakłócenia na poziomie zerowym (równym jego wartości oczekiwanej), co upraszcza formułę, według której prognozujemy

$$\hat{y}_\tau = \hat{\mathbf{x}}_\tau \hat{\mathbf{b}} + 0, \quad \tau > T. \tag{6.13}$$

Jednakże jeśli w równaniu (6.12) zakłócenia wykazują autokorelację – wykorzystanie tej informacji może pozwolić na uściślenie prognozy poprzez wprowadzenie do powyższej formuły przewidywanych wielkości zakłóceń. Obserwacja reszt zrealizowanych w okresach poprzedzających okres prognozy może dostarczyć ważnych wskazówek – w szczególności pojawienie się ciągu reszt o jednakowym znaku może być interpretowane jako sygnał, iż w równaniu zaszła **zmiana strukturalna**.

Zrealizowany błąd prognozy ex ante generowanej za pomocą równania (6.12) zapiszemy następująco:

$$\begin{aligned} \tilde{u}_T &= y_T - \hat{y}_T \\ &= \mathbf{x}_T\mathbf{b} + u_T - \hat{\mathbf{x}}_T\hat{\mathbf{b}} + (\mathbf{x}_T\hat{\mathbf{b}} - \mathbf{x}_T\hat{\mathbf{b}}) \\ &= (\mathbf{x}_T - \hat{\mathbf{x}}_T)\hat{\mathbf{b}} + \mathbf{x}_T(\mathbf{b} - \hat{\mathbf{b}}) + u_T \end{aligned} \tag{6.14}$$

gdzie, dla uwypuklenia źródeł błędu prognozy, po drugim znaku równości dodaliśmy i odjęliśmy wyrażenie $\mathbf{x}_T\hat{\mathbf{b}}$, a po trzecim znaku odpowiednio pogrupowaliśmy składniki wyrażenia. Powyższy wzór pokazuje trzy główne źródła błędu prognozy ex ante. Nieznajomość prawdziwych wartości zmiennych objaśniających $\mathbf{x}_t$ jest pierwszym źródłem błędu prognozy – wykorzystanie ich przybliżonej wartości prowadzi do błędu $(\mathbf{x}_t - \hat{\mathbf{x}}_t)$. Drugie źródło wiąże się z faktem, że nie znamy prawdziwych wartości parametrów $\mathbf{b}$ równania, a jedynie ich oszacowania $\hat{\mathbf{b}}$ wnoszące błąd estymacji $(\mathbf{b} - \hat{\mathbf{b}})$. Ze wzorów dla równania o jednej zmiennej objaśniającej wiemy, że wpływ błędu estymacji narasta w miarę jak wartości zmiennych objaśniających oddalają się od swoich średnich w próbie. Trzeci człon jest związany ze stochastyczną naturą związku zachodzącego pomiędzy zmiennymi równania i jest reprezentowany przez zakłócenie u.

Jak wynika ze wzoru – błędy wynikające z niedoskonałej znajomości wartości zmiennych równania i niedokładnej znajomości jego parametrów splatają się z sobą. Choć możliwe jest, że ich efekty się znoszą, to równie dobrze mogą się wzajemnie wzmacniać.

We wzorze (6.14) nieobecne jest czwarte źródło błędu, pojawiające się wówczas, gdy rzeczywista postać opisywanego związku jest odmienna od tej, którą przyjęliśmy do szacowania, a następnie do prognozowania.

Szczególnym przypadkiem tego błędu jest sytuacja, gdy w okresie objętym próbą wykorzystaną do estymacji postać funkcyjna była poprawna, natomiast w okresie prognozy uległa dezaktualizacji – czy to dlatego, że w okresie prognozy zmianie uległa sama postać funkcji, czy też przez to, że zmieniły się wartości jej parametrów.

W przypadku prognozy ex post, tj. prognozy wykorzystującej rzeczywiste (zrealizowane) wartości $\mathbf{x}_t$ zmiennych objaśniających, błąd prognozy wyraża się następująco:

$$\begin{aligned} \tilde{u}_T &= y_T - \hat{y}_T \\ &= \mathbf{x}_T\mathbf{b} + u_T - \mathbf{x}\hat{\mathbf{b}} \\ &= \mathbf{x}_T(\mathbf{b} - \hat{\mathbf{b}}) + u_T \end{aligned} \tag{6.15}$$

i nie zawiera skutków niedokładnego wyboru wartości zmiennych objaśniających.

Określenie parametrów rozkładu błędów $\mathbf{v}_T = \mathbf{x}_T - \hat{\mathbf{x}}_T$ prognoz zmiennych objaśniających $\mathbf{x}_T$ napotyka na olbrzymie trudności. W praktyce ograniczamy się do wyznaczania wariancji błędu prognozy ex post, zakładającej doskonałą znajomość wartości zmiennych objaśniających $\mathbf{x}_T$. Jeżeli parametry równania